DOCUMENT RESUME

ED 206 427 RC 012 815

TITLE	Supplement to the Final Report on the Edgewood School Plan. Volume II.
INSTITUTION	Development Associates, Inc., Arlington, Va.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Aug 77
GRANT	OEC-0-72-5309
NOTE	72p.; For related documents, see ED 201 450-452 and RC 012 814.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; Affective Behavior; Bilingual Education; Communication Skills; Comparative Analysis; *Curriculum Development; *Educational Change; Elementary Secondary Education; History; Language Dominance; *Mexican American Education; Mexican Americans; Motivation; Program Costs; Program Effectiveness; *Program Evaluation; Reading Achievement; Self Concept; Social Attitudes; Student Attitudes; Student Placement; Teacher Attitudes; Teacher Behavior
IDENTIFIERS	*Edgewood Independent School District TX; *Experimental Schools Program; Purdue Social Attitudes Scales; Texas (San Antonio)

ABSTRACT

Prepared to assist professional educators and researchers who want greater detail on certain aspects of the evaluation of the five-year Experimental Schools Program conducted by the Edgewood Independent School District in San Antonio, Texas, this supplement to the Final Report on the Edgewood School Plan is composed of abstracts that give the basic design and instrumentation of the research, the findings from that particular investigation, and the conclusions that were drawn. The abstracts, which are arranged chronologically by submission date, describe historical background of the Edgewood District; self-concept studies of Mexican American students; first and second interim reports on the program; a comparative analysis of English and Spanish reading performance of Mexican American students; an instrumentation study of the Purdue Social Attitude Scales for Primary Children; social attitudes of South Texas primary students; locus of control and achievement of Mexican American students; a case study of language dominance test development and instructional placement, psychosocial growth and academic achievement of Mexican American students; an instrumentation study of the Self-Concept and Motivation Inventory; an explanation of educational change (case studies from Edgewood); teacher classroom behaviors; the communications skills program; a summary review of revenues and expenditures for the program; program curriculum and instruction; and formative evaluation in the Edgewood School Plan. (CM)

ED206427

D. C. DEVELOPMENT ASSOCIATES, INC.
MANAGEMENT AND GOVERNMENTAL CONSULTANTS
2924 Columbia Pike
Arlington, VA 22204

VOLUME II,
SUPPLEMENT TO THE FINAL REPORT ON THE EDGEWOOD SCHOOL PLAN

Development Associates, Inc.

August 1977

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

RC 012815

The research reported herein was supported by grant OEC-0-72-5309 from the National Institute of Education, Experimental Schools Program, of the US Department of Health, Education, and Welfare. The opinions expressed and the conclusions drawn do not necessarily reflect the position nor policy of the National Institute of Education and no official endorsement should be inferred.

320 Park Drive, San Antonio, Texas 78212 - Phone (512) 824-2724

PREFACE

Volume II, Supplement to the Final Report on the Edgewood School Plan, was prepared to assist professional educators and researchers that want greater detail on certain aspects of the evaluation of this five year experimental program. While these abstracts are by no means complete, they do give the basic design and instrumentation of the research, the findings from that particular investigation, and the conclusions that were drawn. Still further details are available, of course, in the publications themselves. At the time this report was completed, requests for those publications were to be directed to:

Office of Experimental Schools
National Institute of Education
1200 - 19th Street, N. W.
Washington, DC 20208

Although the Institute plans to transmit these for inclusion in the ERIC system, that had not yet been accomplished at the conclusion of the research.

The abstracts have been subjected only to very minor editing even though this resulted in some redundancy among them, primarily because it was expected that not everyone would be interested in all of them. Each thus remained intact enough to provide a sufficient description of the work.

The Development Associates proposals, technical papers prepared for the joint Development Associates/National Institute of Education seminars, and smaller papers submitted to facilitate the consideration of the evaluation design or instrumentation, were not included in this supplement. The essential portions of their content were incorporated into the pertinent reports.

Similarly, although many documents of this type also provide the bibliographies from the reports, they were too voluminous for inclusion. Citations for the instrumentation, important guiding literature, and some Edgewood Independent School District and Edgewood School Plan publications are found in the bibliography to the narrative of the final report. Readers interested in complete documentation are invited to consult the bibliographies of each report.

Three ways were suggested for organizing the abstracts: subject, chronologically by submission, and chronologically by period covered. Because of the breadth of the investigations and because both the content and periods examined were often overlapping, the abstracts were finally arranged chronologically by submission date. The bibliographic citations to the final narrative report direct the reader to the appropriate sections of the individual reports, and therefore to the abstracts, thus facilitating examination by either subject or period examined.

TABLE OF CONTENTS

Section		Page
PREFACE		i
I	A historical baseline of the Edgewood Independent School District. Munoz, Adolfo & Jones, Earl. November 20, 1972.	1
II	First Interim Report on the Edgewood School Plan. Development Associates. March 12, 1974.	3
III	A multidimensional study of self concept of Mexican American students. Sharp, Michael; Cervantes, Robert A. & Jones, Earl. December 1975.	12
IV	Second interim report on the Edgewood School Plan. Cervantes, Robert A. & Jones, Earl. January 1976.	13
V	A comparative analysis of English and Spanish reading performance of Mexican American students. Cervantes, Robert A. & Bernal, Helen H. June 1976.	29
VI	An instrumentation study of the Purdue Social Attitude Scales for Primary Children (English and Spanish versions). Jones, Earl. October 1976.	31
VII	Social attitudes of South Texas primary students. Jones, Earl. November 1976.	33
VIII	Locus of control and achievement of Mexican American pupils. Cervantes, Robert A. & Jones, Earl. December 1976.	35
IX	Language dominance test development and instructional placement: a case study. Cervantes, Robert A.; Theimer, William C. Jr. & Jones, Earl. December 1976.	36
X	Psychosocial growth and academic achievement of Mexican American students. Cervantes, Robert A. & Bernal, Helen H. March 1977.	38
XI	Academic achievement of Mexican American students: the Edgewood School Plan. Welch, Richard O. April 1977.	41
XII	An instrumentation study of the Self-Concept and Motivation Inventory (English and Spanish versions). Jones, Earl. June 1977	44

Section		Page
XIII	The explanation of educational change: case studies in the Edgewood School Plan. Bernal, Helen H.; Jones, Earl & Cervantes, Robert A. June 1977.	49
XIV	Teacher classroom behaviors. Potter, Francis J.; Jones, Earl & Cervantes, Robert A. June 1977.	52
XV	The communications skills program in the Edgewood School Plan: a case study. Cervantes, Robert A. & Jones, Earl. June 1977.	53
XVI	Summary review of revenues and expenditures for the Edgewood Independent School District and the Edgewood School Plan. Potter, Francis J. July 1977.	56
XVII	Curriculum and instruction in the Edgewood School Plan. Jones, Earl. June 1977.	57
XVIII	Formative evaluation in the Edgewood School Plan. Theimer, William C. Jr. August 1977.	62

A HISTORICAL BASELINE OF THE EDGEWOOD INDEPENDENT SCHOOL DISTRICT

The Edgewood Independent School District (EISD) was awarded a grant for an innovative school plan by the National Institute of Education's Experimental School Program in 1972. The district proposed to develop a relevant, compatible program designed to raise pupil achievement through improved administration and instruction coupled with high community/parental involvement. Both internal (formative) and external (summative) evaluations were set up, concomitant to program initiation. This historical baseline report was prepared by the external evaluators to give direction to the coming research.

EISD, located in the City of San Antonio, Bexar County, in upper South Texas, was de facto segregated with 90% of the population Mexican American. It was one of the economically poorest in the state. It's curricula had been traditional, the board only recently modified somewhat from conservative philosophy, and its facilities were, for the most part, badly in need of renovation or reconstruction. The district had tried to force consolidation with the more affluent San Antonio Independent School District, but had failed. A famous legal battle, the Rodriguez case, had been fought into the U.S. Supreme Court to require equalized state funding of schools. The community had also won a court case that stopped federal construction of further low cost housing in the area because of the overcrowded conditions in the schools. Despite these signs of an aroused community, voting turnout generally remained quite low.

Students had recently voiced, through a boycott, great dissatisfaction with their education. Subsequently, some changes were made in curricula, teaching, and administration. The superintendent and later the acting superintendent resigned, and a new one hired. Increased federal funding and further curricular changes occurred. Model Cities aided with building renovation and construction and paid the costs of a bilingual early childhood program.

Through the report period, gradual improvements in property tax assessments occurred (the local percentage share was unchanged), the state's funding raised (but stayed from 66 to 68% of the total), and the federal government's contributions went from about 20% to nearly 23% of the total funds expended by EISD. Per capita costs increased from $380.06 to $538.65 but it still remained one of the lower in the county. Bonded indebtedness and tax rates continued to be among the smallest in the area. Average daily attendance, 91 to 93%, continued to be a problem but was not greatly different from other districts its size. Personnel-pupil ratios were high; teacher salaries were fixed at the state minimum; non-degreed and non-certified teachers made up too large a portion of the staff. Teacher turnover was very high.

One high, one middle, and four elementary school sites were chosen. Staff development, planning activities, and curriculum investigation were begun. The external evaluation cited four primary areas of probable fruitful research: educational change, socio-economic factors, parental

involvement, and fiscal patterns. Negotiation was begun with the National Institute of Education on the evaluation research design, instrumentation, and reporting systems.

FIRST INTERIM REPORT ON THE EDGEWOOD SCHOOL PLAN

The Edgewood School Plan had just completed the first fifteen months of a five year experiment sponsored by the National Institute of Education, Office of Experimental Schools. The first interim report provides enough information to characterize the background against which the project was conceived, the project itself, the external evaluation design, and the principal evaluation findings with a minimum of detail.

Overview

The Edgewood Independent School District (EISD) is a de facto segregated school district located on the west side of San Antonio, Texas. In 1970, the population was 73,000. According to the census taken in that year, fifteen percent of the population over the age of 25 had completed "less than one year" of schooling and sixteen percent had completed high school. Twenty-eight percent of the families in the District had incomes below the poverty level established by the federal government. Over ninety percent of the residents was Mexican American, compared with fifty-two percent for San Antonio as a whole.

For the thirteen years preceding 1968, the district had the same Anglo superintendent who was annually reappointed by an all Mexican American school board made up largely of businessmen. In that year, however, a series of student and community protests over such issues as fiscal policy, teacher qualifications and ethnicity, physical plants, and the recognition of students' rights precipitated a number of rapid changes, resulting in changes in the board's membership and makeup, that was followed by the appointment of a new superintendent of schools.

Corporal punishment was prohibited in the entire district; the number of degreed teachers rose from 81% to 89% between 1968-1969 and 1972-1973, while during the same period the number of Mexican American teachers rose from 28.4% to 36%; and in cooperation with the Model Cities Program an early childhood development center, an elementary school and a middle school were ordered, funded, and built between 1968 and 1972.

There were twenty-six schools serving a student population of 23,000, of which 90% was Mexican American, 6% Black, and 4% Anglo. The basic profile was one of low per-pupil expenditures, low student achievement with a progressively widening gap in grade equivalency scores as measured by standardized testing instruments, high dropout rates, and high in-grade retention rates at the middle and high schools.

The Edgewood School Plan

The Edgewood School Plan (ESP) is part of a nationwide program sponsored by the National Institute of Education, Office of Experimental Schools, in which selected districts have the opportunity to identify local educational needs and problems, and to develop and implement a comprehensive plan that reflects the thinking, solutions, and strategies perceived to

be viable or important by local decision makers. The unique characteristics of this program are the relative autonomy of the local program staff, the long term funding involved, and the long term evaluation effort that forms an integral part of the program.

The Edgewood School Plan, formerly called the Edgewood Experimental Schools Project, was designed in the context of federal guarantees of a five year funding period. It is a program intended to improve student performance through a process of educational change motivated by the "theory of incompatibilities." This theory is presented in the Final Plan, Experimental Schools Project, dated May 1972, and formed the basis on which federal funding was obtained.

Briefly, the theory of incompatibilities asserts that the schools represent a traditional system "designed for White, English speaking, Anglo Saxon, middle class oriented children," staffed by teachers who are trained only to teach such children, and which therefore fails to speak to the "cultural and learning characteristics of the student population it purports to serve." The overall goal of ESP is to redesign the educational program to eliminate existing incompatibilities and to replace them with a program which "reflects the student's cultural, language, and economic characteristics, and is compatible with his learning style." The Final Plan cited five specific areas in which incompatibilities arise in the district: poverty, mobility, culture, language, and societal perceptions.

The Final Plan named six of the twenty-six EISD schools to take part in the Edgewood School Plan: four elementary, one middle, and one high school. Of the total ESP enrollment of 5,367 in 1972, over ninety-six percent was Mexican American, three percent Black, and less than one percent Anglo In September 1972, the starting point for ESP operations, the six schools had a total of 221 teachers (a teacher-student ratio of 24:1), of whom approximately forty percent was Mexican American, nine percent Black, and fifty-one percent Anglo.

At the start of the program, the ESP staff numbered 45, including six counsellors, four instructional coordinators, two scheduling directors, and associate principals for the six treatment schools. The balance of the staff, which numbered about 25, formed the ESP central staff, and served in a resource function for the project. The internal design of the ESP central staff, as outlined in the Final Plan, consisted of a director, and a number of components, each assigned specific categories of tasks such as designing new curricula, providing training for teachers, monitoring products for cultural relevance, providing social services to the community, redesigning the program for educationally handicapped students, providing educational alternatives for students through a program of contracted instruction with non-school personnel, and evaluating both student performance and internal activities.

The External Evaluation

The evaluation of the Edgewood School Plan was keyed to two major constructs. The National Institute of Education, Office of Experimental Schools, has as one of its high priorities, an investigation of the construct of "holistic" or "comprehensive" change. This construct implies the involvement of the total school program of the target population, and the implementation of a systematic, system wide program design, rather than a series of piecemeal changes.

The other construct that structures the external evaluation is that found in the Final Plan in the form of the theory of incompatibilities, Edgewood's interpretation and conceptual application of the construct of "comprehensive change" to the local situation. That is, the entire system of schooling is to be changed in order to eliminate perceived incompatibilities or conflicts between the traditional White, middle class educational system and the characteristics of the Mexican American learner. In order to deal with these two elemental constructs, the external evaluation team has established four principal dimensions of analysis: educational change, educational compatibility, social context, and student impacts.

Educational change was defined as a modification in the school system intended to produce a change in the student's perceptions, attitudes, knowledge, and/or behavior. Although the external evaluation team assumes that attempted modifications will be comprehensive, those that are piecemeal can also be evaluated within the scope of this definition. Educational change is divided into planning aspects and implementation aspects. The ESP staff and project activities are the primary data sources for this dimension of analysis.

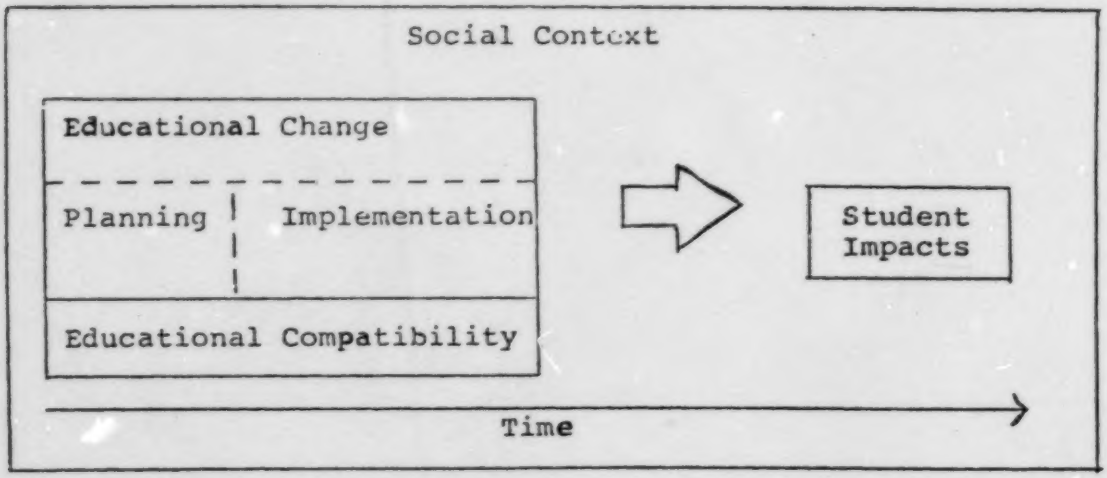
Educational compatibility was defined as a teaching-learning process that is identifiable in an educational system in which the experiences, needs, and learning characteristics of ESP students are recognized and productively incorporated into essential personal relationships and perceptions, curricular materials, methodological approaches, and school policies. This definition incorporated the assumptions of the Final Plan, i.e., that the primary sources of incompatibilities were socio-cultural and economic. It moved beyond this assumption, however, by additionally incorporating as sources of incompatibilities, broader pedagogical and methodological considerations involving the learning characteristics of individual students. Teachers and principals were the primary data sources for this dimension of analysis.

Social context was defined as the relationship of the program with various external intervening variables which could facilitate or hinder the progress of the program. The external variables may be initiated in the district, the state capital, or in Washington, D.C. Parents and the school board, however, are the primary data sources for this dimension of analysis.

Student impact was defined as the variables which will measure student progress associated with program treatments intended to make the educational system more compatible. While it is recognized that the project

can have measurable impact on parents, teachers, principals, or other groups, the external evaluation design specified the dimension of student impacts as the ultimate dependent variable.

Incorporating the dimensions of analysis into an evaluation design, educational change and educational compatibility were considered as project means, treatments, or (in terms of research design) independent variables. Student impact was considered the project ends, effects, or dependent variables. The dimension of social context was viewed as consisting of a number of intervening variables. Diagramatically, the evaluation design appears as follows:



Given a comprehensive project involving multiple change being introduced simultaneously, and an equally comprehensive evaluation design, the "laboratory" determination of unequivocal cause and effect is visibly impossible. The external evaluation design deals with the matter in two ways. Primarily, treatments and impacts will be subjected to a two-stage analysis. In the first, treatments are measured along with impacts. In the second, each treatment will be associated either logically or statistically with its most probable impact or impacts. Cause and effect can then be inferred from a strong association between treatment and impact.

The other way the design helps to imply cause and effect relationships is through the use of comparison groups. In addition to tracking treatments and impacts in the six ESP treatment schools, the same data from students in four comparison schools are collected. If a change occurs in the data gathered in the treatment schools, and no change occurs in the same data from the comparison schools, it is plausible that the former change can be associated with some project treatment.

Educational Change

The dimension of educational change involved essentially a documentation function, and placed maximum emphasis on the criterion of comprehensiveness. The changes described under this dimension were to be analyzed

similar to that for the dimension of educational compatibility.

The initial planning phase of the ESP that led to the production of the Final Plan was weak in three major areas. First, school site personnel, students, and community representatives were neither consulted nor involved in the initial planning or definition of the project. Second, no formal needs assessment was performed for any potential target group. Third, perhaps as a consequence, the Final Plan was stated in fairly general terms, particularly in the areas of behavioral, attitudinal, instructional, and policy changes anticipated or recommended.

Program implementation is an important object of study since it is assumed that the quality or effectiveness of the implementation stage will be closely associated with the quality of the impacts. The initial implementation of the program was hampered by a number of factors.

First, there was little opportunity to justify and explain the project to the school personnel who would ultimately be responsible for its operational success or failure. Second, since funding was delayed, there was little pre-implementation training for ESP personnel. Third, the new roles and functions of ESP staff caused some ambiguity and concern on the part of ESP principals, who stated that they did not know some of the ESP people who were coming into their schools and did not understand their function. Fourth, weak internal communication and coordination also hampered intital project implementation. During the first year, few meetings involving the full ESP staff were held, and even those that involved the ESP director and component heads were suspended between June and September 1973, although the project continued throughout the summer. Fifth, project implementation and subsequent revision necessitated an internal evaluation component.

There was a high turnover of personnel at all project levels. By June 1973, the project had had three directors; five of the six ESP principals had either resigned, requested transfers out of ESP, or been fired; eleven central staff had resigned, as had nine ESP school site personnel. This represented a turnover rate of 57% among ESP central staff and 63% among school site personnel. Data from the district as a whole were not available.

In September 1973, early into the second program year, the position of internal evaluation director was staffed. Planning was made the highest priority immediately. In October, a preliminary needs assessment was conducted and a four day planning retreat for ESP staff and principals was planned for November. As a result of this retreat, the previous staff and school personnel problems diminished. ESP staff began meeting three days a week for intensive planning, and assignments were by "Task Groups," made up of central staff, school site personnel, and principals.

In the area of curriculum, six sets of materials were prepared or sponsored by ESP during the first program year. One of these, Man, A Course of Study (MACOS), was prepared commercially. It was implemented in the middle and high school under the supervision of ESP curriculum component staff. Most other new curricula were locally designed. In cooperation with Title VII, two social studies courses were produced in Spanish for

use in the third and fourth grade bilingual program. Three others were environmental geography, community and local power, and career education, all one semester courses designed around a best paced approach using criterion referenced testing techniques. All of these were offered again during the second program year except for MACOS, which was dropped at the high school when three of the four teachers trained to use it resigned over the summer. One additional course called college bound English was also designed locally and introduced at the start of the second program year. These changes in curriculum were all grade specific.

The most comprehensive of the ESP projects for the first program year, both in terms of process and product, was the best paced mathematics curriculum. The need for full revision of the then existing math program was recognized and voiced by the ESP principals. They carried their requests and recommendations to the ESP central staff. The staff in turn provided manpower, consultants, materials, and teachers for an extensive program, which resulted in a new math program to be coordinated from grades K through 12. It was ready for a pilot operation in the fall of 1973, involving some 3,000 students and 25 teachers.

The major change in the area of participant roles involved the emergence of the principals as instructional leaders. The role implies an increase in the development and implementation of instructional programs and the delegation of some other functions to other personnel.

Educational Compatibility

As to project activities during the first fifteen months, little explicit attention was paid to the concept in terms of large scale operational changes. The breadth of the external evaluation definition of educational compatibility allowed for some observation and baseline data gathering, and a number of individual events occurred relating to compatibility.

Teacher attitudes and perceptions of the organizational climate were measured on the assumption that project treatments may well be aimed at principals and teachers in a way that could produce changes in both, and that other things being equal, schools in which the typical teachers are satisfied and comfortable with the organizational climate will tend to favor teacher motivation which in turn promotes the creation of a more educationally compatible school system.

Teachers' attitudes toward their working environment were not very positive. They reported that morale was low and complained of busy work. Socially, they tended to interact little with their colleagues, finding their best friends elsewhere. Teachers perceived their principal as a neutral factor in their jobs. The majority of the teachers perceived the schools' climate in which they worked as "closed" or "controlling," although in general, elementary schools were perceived as less closed than the middle and high schools.

Teacher-student classroom interactions were measured on the assumption that positive interactions, both verbal and non-verbal, would tend to produce a more educationally compatible system. Working with a modified

observational schedule (OScAR 2a), measures of emotional climate and teacher interpersonal style were calculated. Findings show significantly more positive factors on these measures than negative in all classrooms observed, and that there was no difference between ESP and comparison classrooms. While all teachers observed were more positive than negative, Mexican American teachers showed a significantly higher proportion of positive to negative behaviors than Anglo teachers.

Teacher turnover rates in treatment and comparison schools were obtained on the assumption that, other things being equal, a relatively more stable faculty population would facilitate program implementation. Of a total teaching staff of 221 in the six ESP schools, seventy-five (33.9%) either resigned or transferred. The four comparison schools averaged a 24.3% turnover. For the 1972-1973 school year, then, the turnover rate for the ten treatment and comparison schools combined was 29.1%. Comparable figures for the EISD as a whole were not available.

Social Context

Various factors external to and independent of ESP can at times impinge, positively or negatively, on project planning and implementation. Few board policies had specific effect on ESP during the first fifteen months. A major action affecting the entire district, was when the board transmitted the news that PL874 Federal Impact funds had been reduced by 80% in the district, with the result that all teacher salaries were immediately reduced by $250. This had a predictably negative effect on morale, although the funds were partially restored later in the year. Interviews with four of the board members showed varied opinions about the first period of ESP work and they appeared to be generally positive about its future. Two of the four were not in accord with the theory of incompatibilities, the guiding principle for the project.

National Institute of Education interventions during the first program year were largely limited to providing technical assistance to the project. In late summer 1973, a new program officer was named and the presence of the new project officer at the November planning retreat was cited as a valuable contribution by several ESP staff members.

Student Impact

Baseline data in the area of student impact dealt with self concept, motivation, achievement, perception of school climate, attendance, and dropout rates. Self concept and motivation measures were obtained using the Self-Concept and Motivation Inventory (SCAMIn) on classroom samples taken from kindergarten, first, fifth, eighth, and tenth grades. Data gathered during the spring of 1972 suggested that students typically entered kindergarten with near norm self concept scores and low motivation scores. By fifth grade, however, both self concept and motivation scores were equal to or slightly above instrument norms. Eighth and tenth grades showed the self concept scores remaining fairly e at

or above national norms, but motivation dropped below the norm. The data from treatment and comparison schools found very few statistical differences on any SCAMIn subtests and the two groups were essentially alike.

Achievement data were gathered on the same grade samples (K, first, fifth, eighth, and tenth) used for the SCAMIn. The test instruments used were the Tests of Basic Experiences (TOBE) for grades K and one, and the Comprehensive Tests of Basic Skills (CTBS/Q) for grades five, eight, and ten. Two interesting facts emerged from the achievement data. First, that treatment and comparison schools were shown to be virtually equivalent, and second, that the achievement pattern seemed to parallel more closely the motivation level of the classes as measured by the SCAMIn, rather than the self concept level.

Data on student perceptions were gathered using the Student Behavior Description Questionnaire (SBDQ). Grades five, eight, and ten were included in the sample. The treatment and comparison schools were virtually alike. In broadest terms, when compared with national norms, students at all levels measured an average interpersonal need for achievement and task accomplishment, an average interpersonal need for peer affiliation, and less than average feelings of inclusion. Subtest results showed that the typical student felt that his parents showed little support for or interest in his school activities. Teachers were typically perceived as friendly and able to motivate but also as authoritarian and critical. Despite this relatively favorable judgment of teachers, the results indicated little desire to remain in school with low interest in staying out of trouble, accepting school rules, or getting good grades.

During the first program year the combined average daily attendance for the six treatment schools was 92.27%, ranging from a low of 89.75% for the new middle school to a high of 95.63% for one elementary school. In the comparison schools, the average daily attendance for 1971-1972 was 91.78%. For 1972-1973 it was 91.43%, ranging from a low of 89.11% at the high school to a high of 93.04% at one elementary school.

Dropout rates at the elementary level were insignificant. For the two middle schools, the 1972-1973 figures were virtually identical: ESP 1.67%; comparison 1.66%. For the two high schools, however, the dropout rates were dissimilar. For 1971-1972, the dropout rates for the ESP and comparison high schools were, respectively, 7.4% and 4.8%. For 1972-1973, they were, respectively, 6.0% and 4.8%.

Recapitulation

The baseline data gathered reflected a situation of low motivation and student achievement, roughly norm self concept, and some problems with faculty morale. Board member opinions about ESP were mixed but not generally unfavorable. ESP staff changes and additions, plus intensive

needs assessment and planning work, seemed to show a revitalization of effort after a period of uncertainty and some confusion of roles. Incorporation of the principals into the ESP system as instructional leaders was viewed as a positive step. The production of curricular materials was heavy in the early part of the fifteen month period and was to be resumed when the planning process was completed. Staff development had received approval of its plans, for the first time since inception of the project. The internal evaluation component was in operation, also for the first time. The general climate appeared to be improving for the Edgewood School Plan.

A MULTIDIMENSIONAL STUDY OF SELF CONCEPT OF MEXICAN AMERICAN STUDENTS

The purpose of this multidimensional pilot study was to investigate the impact of the Edgewood School Plan on the self concept of a predominantly Mexican American student population. The Edgewood School Plan, Edgewood Independent School District, is a federally funded experimental schools program intended to produce educational compatibility between the educational institution and the needs of a predominantly Mexican American student body.

This study presents a review of the salient literature including definitions of self concept, a survey of studies examining the self concept of Mexican Americans, and the relationship of self concept to achievement. The Self-Concept and Motivation Inventory was the instrument used to measure self concept of students in designated treatment schools as well as comparison schools.

A grade by grade comparison of the 1975 self concept scores, converted to stanines, showed the self concept levels of the third, fourth, and fifth grades in ESP schools to be one stanine higher than the same comparison school grades. The kindergarten comparison group scored one stanine higher than the same ESP grade. All other grades had the same stanine for treatment and comparison schools.

Using a 2 by 2 factorial design, the fall 1974 self concept scores were compared to those from the spring 1975, for treatment and comparison kindergarten students. No statistically significant difference was found between treatment and comparison schools; the students scored higher in spring than when they entered in fall.

Using the 1975 self concept data, the effects of school type (treatment or comparison) and sex were investigated. Using a 2 by 2 factorial design, no significant difference was found between the overall scores of treatment and comparison students. However, when the dimension of sex was considered, females showed significantly higher self concept scores than males.

The Pearson product moment correlation of coefficient was used to determine the relationship between self concept and achievement. The relationship between self concept and achievement was positive but the correlation coefficients were low.

A two factor mixed design with repeated measures on one factor was used in the longitudinal study, conducted to measure the effects of treatment on the self concept of early elementary children. The study tracked students from the first grade, spring 1973, to the second grade, spring 1974, through the third grade, spring 1975. The scores were compared year-to-year, treatment versus comparison, and interaction effects were computed. No difference was found between the treatment and comparison school means. However, the students scored significantly higher in the second grade (spring 1974) than they did as third graders (spring 1975).

SECOND INTERIM REPORT ON THE EDGEWOOD SCHOOL PLAN

The Experimental Schools Program was established by legislation in mid-1970 as a small research program in the U.S. Office of Education of the Department of Health, Education, and Welfare. The operational premise of experimental schools was that educational reform could possibly be enhanced by the long term (five years) support of pilot programs which demonstrated the efficacy of a "comprehensive change" strategy as a means of fostering educational change.

On July 1, 1972, the National Institute of Education was created in PL 92-318 "...to provide leadership in the conduct and support of scientific inquiry into the educational process." Incorporated as a major and vital program in the newly established institute was the Experimental Schools Program (ESP), transferred from the Office of Education. The Edgewood School Plan (ESP), of the Edgewood Independent School District (EISD), San Antonio, Texas, was one of the transferred programs.

In order to assess the efficacy of local attempts at producing educational change, the Office of Experimental Schools (OES) utilized a "levels" concept of evaluation. The first effort, referred to as "Level I," was an internal unit within local programs, vested with a formative evaluation role. The second effort, "Level II," was also site specific but external to the local programs. It was given a summative role carried out by an independent contractor. A "Level III" national cross-site evaluation effort to assess common themes and unique characteristics among all projects had not been operationalized at the time of this report.

Edgewood School Plan

The Edgewood Independent School District, founded in 1913 from two small rural districts, evolved from an Anglo-German farm population to a 90% Mexican American section of San Antonio with a population of 73,000. An all Mexican American board and Anglo superintendent operated, by 1968, a 26 school traditionally oriented district. A student boycott and parent unrest brought board membership changes and later, a Mexican American superintendent. Some management and curricular modifications were also effected. Out of this new direction grew the experimental program in the district.

The Edgewood School Plan, formerly called the Edgewood Experimental Schools Plan, was designed as a program intended to improve student performance through a process of educational change motivated by the "theory of incompatibilities." That theory was presented in the Final Plan, Experimental Schools Project (hereafter, the Final Plan) dated May 1972. Briefly, the theory of incompatibilities asserted that the Edgewood schools represented a traditional system "designed for White, English-speaking, Anglo-Saxon, middle-class-oriented children," staffed by teachers trained only to teach such children, and which therefore failed to speak to the "cultural and learning characteristics of the student population it purported to serve." The overall goal of ESP was to redesign the

educational program to eliminate existing incompatibilities and to replace them with a program that would "reflect the student's cultural, language, and economic characteristics, and be compatible with his learning style." The Final Plan cited five specific areas in which incompatibilities arose in the district: poverty, mobility, culture, language, and societal perceptions.

At the start of the ESP program, the staff numbered 45, including six counsellors, four instructional coordinators, two scheduling officers, and associate principals for six treatment school sites (4 elementary, 1 middle, and 1 secondary). The rest of the personnel, about 25, formed the so called ESP central staff and served in a resource function for the project. The internal operational design of the ESP central staff, as outlined in the Final Plan, consisted of a project director, and a number of "components," each assigned specific categories of tasks such as designing new curricula, providing training for teachers, monitoring products for cultural relevance, providing social services to the community, redesigning the program for educationally handicapped students, providing educational alternatives for students through a program of contracted instruction with non-school personnel, and evaluating both student performance and program-internal activities. Since its inception, ESP had undergone major modifications in organization, staffing, and programmatic thrust.

External Evaluation

The evaluation of the Edgewood School Plan was keyed to two major constructs. The National Institute of Education, Office of Experimental Schools, announced as one of its high priorities an investigation of the construct of "holistic" or "comprehensive" change. This construct implied the involvement of the total school program of the target population, and the implementation of a systematic, district wide program design, rather than a series of piecemeal changes.

The other construct of the external evaluation was that found in the Final Plan in the form of the theory of incompatibilities. The Final Plan was Edgewood's interpretation and theoretical application of the construct of "comprehensive change" to the local situation. That is, the entire system of schooling was to be changed in order to eliminate perceived incompatibilities or conflicts between the traditional White, middle class educational system and the characteristics of the Mexican American learner. In order to deal with these key constructs, the external evaluation team established four principal dimensions of analysis: educational change, educational compatibility, social context, and student impacts.

EVALUATION METHODOLOGY

Educational change was viewed as a process intended to produce improvements in students' affective, behavioral, and cognitive skills using program treatments that departed from conventional Edgewood practices. Educational change was assessed by examining three major areas: treatment documentation,

factors which facilitated or impeded change, and the nature of Level I evaluation.

Evaluation Design

The evaluation design employed in this study was time series and quasi-experimental. The use of this design in the applied educational setting approximated partial control of relevant variables with certain qualifying exceptions and limitations. Within the framework of this design the four dimensions of analysis, educational change, educational compatibility, social context, and student impacts formed the variables for this period of research.

To assess treatment impacts over the intended five year duration of ESP, data were obtained each year as a pre-treatment measure of activities, and each succeeding year as a post-treatment measure. Data were collected each year from seven major reference groups: students, teachers, principals, ESP central staff, EISD Board of Education members, EISD administrators, and parents. Wherever possible, comparison reference groups were used. The data collected from all the groups were synthesized periodically to evaluate the planning process, implementation, and impact of ESP treatments.

This design was intended to be a comprehensive evaluation of the ESP program in a dynamic setting. The design would capture the numerous project treatments attempted, along with influences external to but impinging upon the project, occurring simultaneously. Such a design was more appropriate than one used to assess a "laboratory experiment" in which the total environment would remain constant and one variable at a time would be altered in order to determine its effect or impact on the target population.

To determine cause and effect between treatments and impacts, data were subjected to a two stage analysis. In the first stage, the Level II team documented and measured treatments A_1, A_2, A_3, ...A_n, together with impacts B_1, B_2, B_3, ...B_n. In the second stage, each treatment was to be associated either logically or statistically with its more probably impact or impacts. Cause and effect, then, would be inferred from a strong association between treatment and impact, analogous to an antecedent consequence relationship.

Student Sampling Design

Initially, Level II designated a five year longitudinal tracking system for collecting data on students, the major reference group to be investigated. In the first project year, because of the initiation stage of ESP treatment specification and implementation, Level II tested all students in grades K, 1, 5, 8, and 10 in the ESP and comparison schools. In keeping with the concept of longitudinal tracking of the same students over time, adopted in year one, students in grades K, 2, 6, 9, and 11 were tested in the ESP and comparison schools in the second year of the program.

In the third project year, ESP treatments were implemented and designed to cover all students in grades K, 1, and 2, one-half of the students in grades 7 and 8, and 150 students in grade 9, and to expand the treatments to encompass grades K through 12 by 1977. In order to collect the baseline data needed on students in these grades, Level II conducted supplemental testing in the fall of 1974. Starting in the spring of 1975, annual testing will cover all grades, K through 12, in the treatment and comparison schools. This will provide both post-treatment test data for the second year treatment grades and pre-treatment data for grades to be added to the ESP treatment program the following year.

Statistical Analyses of Data

Each set of evaluation data collected under educational change, educational compatibility, social context, and student impacts was subjected to appropriate statistical treatments to determine discriminatory relationships among the factors or variables and/or the significance of the findings. These methods ranged from the use of descriptive statistics to the more sophisticated techniques involving quasi-experimental designs. Numerous statistical tests were employed such as the chi square, analysis of variance (ANOVA), analysis of covariance (ANACOVA), factorial ANOVAs, correlation and regression analyses, and trend analysis. Assumptions of normality of distribution, homogeneity of variance, and independence of observations were checked whenever appropriate.

EDUCATIONAL CHANGE

The Edgewood School Plan, in its original documentation and subsequent revisions, set out to improve students' achievement in the educational system and eventually the quality of life in the district. It based its methodology on comprehensive change, experimentation compatible with the characteristics of the predominantly Mexican American, low income population. Many modifications were initiated and while they could not be said to represent comprehensive change, they attempted to alleviate some of the students' problems and capitalize on their strengths.

Students comprised the primary target group. To reach them more effectively, the administration, teachers, and parents needed some assistance. These became means to accomplishing the improvement of students. Since day to day labors nearly precluded the formidable task of comprehensive change by those directly involved, an intervention was contrived - the addition of a change agency, the Edgewood School Plan staff. Within this agency specialized functions were set up for the staff members through a component arrangement.

A family services component was instituted to involve parents more deeply in the decision making of the schools and to enlist their support for increased pupil participation in the educational activities. Additionally, during 1973-1974, the component worked directly on providing certain enhancements that were expected to keep absentees and dropouts bodily in the school setting, enabling exposure to the academic treatments. The

component worked with the community at large and parents in particular to increase their participation. There was some evidence of modest success in this realm. The component also provided a wide variety of pupil directed services and while they perhaps resulted in only minor changes in attendance and dropout rates, those services were considered quite useful by the youngsters.

The curriculum component wrote some new materials, revised others to better meet the students' needs, and brought in some programs and methods proven elsewhere. While some of these were small and not measurable within the achievement context, others were broader in scope and deeper in approach. The mathematics and portions of the language programs appeared to increase achievement. Individualizing instruction, best pacing, received a great deal of attention and appeared to be giving results.

Staff development was conducted by a special unit and whether by design, because of staff and teacher turnover, or because of the desire to change the program, primarily repeated the processes of orientation, justification, and modification of the Edgewood School Plan. Teachers received some treatment implementation training but the evidence suggested a great deal more was needed.

Internal evaluation was entrusted to a group called Level I, which spent most of its time conducting a needs assessment, administering a teacher opinionnaire, and exhaustively preparing and administering a language dominance test. Standardized testing was also conducted. This component did not monitor treatments, provide interim testing, nor give input for the improvement of treatments.

A primary function of any organization, planning, was not delegated to a separate authority but was included in the tasks of all ESP staff. Characterized as inadequate in the first year of operation, planning was given top priority in 1973-1974 and it dominated the experimental process. Detailed process description was provided in Development Associates' first interim report, which covered half of the planning year. Two major consequences were determined. The first was that almost all other activities were relegated to what was termed maintenance status. The programs most affected by this holding action were the classroom treatments and the training of the implementing teachers. Nearly nothing was done to perpetuate these and many dissolved due to insufficient attention.

The second result was weak planning. Since the basic elements - compatible educational changes - were either never defined or the definition (especially as translated into action) was rejected, planning tended to vacillate between clinging to those decisions begun out of the Final Plan or inventing new ones seen as revolutionary, or at least more in step with the ideas of the ESP staff. A part of this difficulty lay in the concept of compatibility, or perhaps incompatibility, stressed in the documentation. Development Associates' efforts to research the characteristics of the population showed a moderately conservative orientation of the student body, community, and teaching faculty, not greatly different from those found anywhere. The strategy had proposed the creation of totally new solutions to the problems. The resulting confusion, epitomized in the planning process, rendered comprehensive change almost unattainable.

EDUCATIONAL COMPATIBILITY

A fundamental premise of the Edgewood School Plan was the theory of incompatibilities, which asserted that the traditional educational system was incongruent with the needs, experiences, and learning characteristics of low income, Spanish speaking, Mexican American children. A principal goal of ESP was the development of an instructional program that was consonant with and responsive to the needs of these children. The concept of compatibility represented a major theoretical guide to the planning and implementation of change.

The Edgewood Final Plan definition of educational compatibility was stated more in terms of a problem rather than a systematic theory. A series of symptoms suggesting the problems were delineated. The ESP staff, in a Communications Skills Philosophy paper, attempted to operationalize the construct; however, its characteristics and parameters remained somewhat ambiguous.

Level II defined educational compatibility in general terms in the First Interim Report. In consort with its technical panel and the NIE/OES project officer, the definition was reformulated as follows: a flexible teaching-learning process in which experiences, needs, and learning characteristics of the student are purposefully the focus of personal relationships and interactions, curricular materials, methodological considerations, and educational policies; the end result would be optimal student learning. This definition was essentially pedagogical and recognized the need for congruence between the instructional/learning process and the experimental characteristics of the pupil.

Six variables were examined within the dimension of educational compatibility:

- School Organizational Climate
- Teacher Perceived Patterns of Pupil Control
- Teaching Ambiance
- Teacher Turnover
- Teacher-Student Values
- Locus of Control

Primary emphasis was placed on teacher variables, based on the premise that teachers represented a critical factor in the instructional process and ESP. The ultimate success of the Edgewood School Plan will, to a large degree, be predicated on the ability of teachers to carry out the program.

The study of organizational climate indicated that the ESP schools were "closed." According to previous studies, a closed climate was not conducive to innovation or change. Teaching ambiance was essentially traditional. Teacher turnover rates were high. Previous studies had associated traditional, closed environments with high teacher turnover. Since both the ESP and comparison schools were closed, it appeared that ESP had so far been unable to make changes in the general working climate for teachers.

Within a school system which was characterized by a closed and traditional structure, pupil control, as perceived by teachers, was found to be moderate. This was particularly true of pupil access to materials and general mobility with the classroom.

During thissame reporting period, ESP treatments were implemented and some positive differential effects were found in student achievement. It appeared that treatment teachers had created and maintained moderately open sub-cultures in their classrooms, which were amenable to the introduction of innovation. These classrooms then represented fairly open subsystems functioning within the closed/traditional ethos of the school. Teachers conformed to the norms of the school system, but at the same time established less rigid systems in their respective classrooms, allowing students greater freedom of movement and autonomy.

The studies of teacher-student values, and teacher locus of control began to explore those factors directly affecting the teaching-learning process. The continued study of the compatibility factors is expected to provide insight into the characteristics of the environment conducive to innovation.

SOCIAL CONTEXT

The dimension of social context in this investigation was a conceptual application of the social system framework used in the investigation of educational systems. Specifically, the dimension of social context referred to selected external intervening variables which were distinct but which impinged on the operational activities of the Edgewood School Plan:

- EISD Board of Education in terms of the members' perceptions and understandings of ESP, and their relationships with the community and the superintendent;
- EISD board policies related to ESP;
- EISD funding variations within operational funds and ESP budget expenditures.

Some degrees of community school interaction were probed in board and other election voting patterns, trends in school attendance, and other commitment indicators. The relationships between these and ESP goals and operations were expected to predicate program trends.

The research in the social context impingements on the ESP change process led to the conclusions that:

1. The EISD board was, as in the previous year, only vaguely familiar with the purpose, goals, and objectives of ESP and did not appear to completely share the assumptions of the ESP Final Plan or the present ESP operation, nor did it exhibit extraordinary support of ESP.

2. The EISD operational funding had remained relatively the same as the previous year, apparently reflecting a local decision to cut back on programs.

3. ESP budget expenditures tended to reflect poor accounting procedures and weak program planning and management.

It would appear the board, possibly because of its moderately conservative nature and lack of substantive involvement in ESP, had contributed to a painfully slow process toward meeting the ESP goals and objectives.

Although the district faced serious financial problems as did other schools in South Texas, its operational budget had remained relatively stable, being affected mostly by inflation. Federal funding for ESP, or potential lack of it, did not at this point appear to have significantly altered the fiscal status of EISD.

There did not appear to exist a clearly defined procedure for allocation of ESP budget expenditures and budget accounting procedures were found to be weak. This appeared to reflect a lack of substantive planning and management procedures in relation to defined goals and objectives. To maximize dollar impact on fostering educational change, more appropriate procedures were warranted.

The impact of the researched social context factors on the educational process appeared to result more from default than conscious influence. Most of the board members knew little about the Edgewood School Plan and were not actively pursuing familiarity with it. The administration seemed to reflect this general board inaction. The positive changes found in impacts, then, were occurring in spite of the researched social context factors, not because of them.

AFFECTIVE IMPACTS

Affective variables were incorporated into the study of the Edgewood School Plan for two very important reasons: affect is a primary factor in the amount of achievement realized, and, the founders specifically noted that pupils' attitudes toward education were inadequate and needed to be changed. The move toward compatibility addressed the first of these reasons. The idea of comprehensive change was expressed with the expectation that greater success in school would improve general attitudes and result in increased persistence in school and active support of education by both students and the populace.

Narrowing the impossibly broad studies of total self concept, the study selected school self concept as a manageable variable. It, along with motivation, was examined. Attitudes, also restricted to those connected with school, were measured. Comparisons among students for the first two years were available for this study; others will be available during succeeding years. Eventually, the relationships to other student variables, achievement and school persistence, will be sought.

The investigation of school environment, self concept and motivation, stemmed from the need to assess the validity of two statements made in the Edgewood Final Plan. The designers of that document stated that "children often come to school with negative feelings about self" and "...collision with school usually reinforces these feelings or creates new ones." These two school variables were assessed with the Self-Concept and Motivation Inventory (SCAMIn). The SCAMIn measured students' concepts of themselves in certain areas related only to school. It pretended no measurement of general self concept nor motivation, nor to any specific area except that connected with school.

All students in grades K, 1, 5, 8, and 10 in the ESP schools (N=1,442) and approximately 120 students in each of the identical grade levels in the comparison schools (N=627) were tested with the SCAMIn during the first project year. This test administration was part of a five year longitudinal study designed to collect affective, cognitive, and behavioral measures.

In keeping with the concept of longitudinal testing of the same students over time, all students in K, 2, 6, 9, and 11 in the ESP and comparison schools were administered the SCAMIn during the second project year. Thus, all 1,574 students in grades K, 2, 6, 9, and 11 in four elementary and one high school designated as ESP schools, and all 1,145 students in the identical grades in two elementary and one high school designated as comparison schools, were administered the SCAMIn. A combined total of 2,719 students were tested.

Data from the first two years of study indicated that in self concept:

- kindergarten students scored slightly below the norm,
- first graders scored equal to the norm, and
- second grade students were above the norm in self concept.

Thus an upward trend in self concept, across the primary grades (K→1→2), was suggested.

Data on other closely related Mexican American groups were scarce, especially those utilizing the SCAMIn instrument. Raw data were obtained from a much smaller (1528 average daily attendance) South Texas school district with a 98% Mexican American population. Stanines identical to those calculated for Edgewood were reported for grades 1, 2, and 5.

The present study showed that for Edgewood, the self concept in grades 8 through 11 was relatively stable when expressed in stanines. Grades 8, 9, 10, and 11, tested in 1973 and 1974, scored equal to the norm. The results of this study, supported by the Cicirelli and South Texas research, suggested an early rise, then fluctuation in self concept in the elementary school. Studies in middle and secondary level self concept

were either difficult to compare or inconclusive. This study, however, did point to a relatively stable self concept for middle and secondary school level students.

Student Attitudes

Students' attitudes towards their school environment is an important indication of school climate and efforts to bring forth positive changes in the educational process. To measure students' attitudes and perceptions of their schools, teachers, course work, and social activities, Level II utilized the Student Behavior Description Questionnaire (SBDQ).

The results of the SBDQ administration and analysis led to these conclusions:

- The Edgewood students, both in ESP and in comparison schools, demonstrated attitudes and perceptions concerning their school environment which were very similar to the responses of students in other schools.

- The classification of a large proportion of Edgewood students as ambivalent in student type was also typical of the responses of other studied student groups.

The findings that Edgewood students' attitudes toward school were similar to those of most other groups was not expected in light of the literature. The studies suggested that Mexican American students' value orientation and attitudes were different from those of their Anglo counterparts.

The Edgewood study of student attitudes not only failed to confirm the literature, it found contrary conditions:

- Edgewood students' attitudes were about like those of any other US group they formed a decided cluster around the norm.

- The highest frequency student type was "ambivalent," that too, was the norm for most schools.

Furthermore, when this study was added to the findings on self concept and motivation, corollary results occurred:

- Middle and high school students' self concept was mostly at the national norm.

- Those same groups' motivation also scored near the norm.

Additionally, the exploratory student values study discovered:

- Middle school students, whether of low or middle income status, subscribed to what were thought to be middle class values.

- They tended to be generally toward the conservative side in their school value orientations.

These sets of data and their conclusions may be of critical importance. These students were measured, not by a single instrument and not in one single time frame; three instruments and two years' study on two of them were included.

The implications are far reaching. ESP had experienced considerable difficulty with some of its innovative programs. It seemed possible that some of these may have been incompatible with the middle of the road attitude set of the Edgewood students, thus causing disorientation and perhaps even rejection.

BEHAVIORAL IMPACTS

The ESP Final Plan cited poor attendance, high in grade retention rates, and high dropout rates in the Edgewood school district as indicators of the need for changed and improved educational programs. The evaluation study, taking its cue from that report, limited its behavioral variables very narrowly, precisely to these three elements. They were examined by Level II because of their stated importance to ESP and because it was assumed that they would provide objective measures of an important aspect of the experimental environment not addressed by other measures. It was hypothesized that in 1973-1974, if positive ESP treatments were implemented, they would lead to increased attendance, lower in grade retentions, and lower dropout rates.

To test the hypothesis, data for these variables were collected from ESP and from comparison schools in the first and second years. The availability of these data enabled the comparison of the variables from 1973-1974 to the baseline year (1972-1973), and a comparison between the treatment and comparison schools' variables within the same year period. All information concerning the behavioral variables was derived from the Edgewood school district administration and/or the principals' term reports.

The behavioral data led to the following conclusions:

- ESP attendance for the second year of the project decreased from the first year attendance rate.
- ESP and comparison attendance were similar.
- Overall ESP in grade retention rates increased for 1973-1974.
- The ESP middle school showed decreased in grade retention.
- Overall ESP dropout rates increased for 1973-1974.
- The ESP high school showed a slight decrease in dropouts.

- The presence of ESP sponsored programs had no positive effect on attendance, on in grade retentions in the elementary schools or the high school, or upon dropout rates in the middle school.

- The presence of ESP sponsored programs appeare to have had a positive effect on in grade retentions ... the middle school and may have had a positive effect on high school dropouts. Looking at all the behavioral variables as a whole, the data did not show a positive overall change.

The ESP treatments which had the potential for affecting the behavioral variables were a truancy reduction program, and changed curricula and instructional methodology. The truancy program dealt directly with the three behavioral variables of interest, for reduced truancy would improve attendance and thus lead to reduced in grade retentions and school dropouts. The family services component of ESP, which directed the program, contacted and attempted to assist truants and/or potential dropouts so that they could return to school.

Because the curricular and methodological changes were believed to be those which could bring about more student enjoyment and achievement in school, they could be assumed to also cause better attendance, less in grade retention, and fewer dropouts. The closest relationship would appear to have been the ESP high school where achievement scores were higher and where there was a slight reduction in dropouts. Greater satisfaction could have kept some of the students in school.

The 1973-1974 data did not indicate, to the degree hypothesized, that the treatments had led to the general behavioral consequences for which they were intended. There were two possible explanations for this finding. Either the ESP treatments were not as effective as they should have been on the low attending, potential dropout student, or the evaluation time period was too short for such changes to have occurred, or both. Other research has found that only on going, long term programs which began with kindergarten children had an effect on these behavioral variables. Such a program's effect, the studies suggested, could only be fairly evaluated after many years.

COGNITIVE IMPACTS

A primary indicator of program impacts was the examination of student cognitive variables, defined by Level II in terms of student performance in reading, general language, science, social studies, arithmetic, and basic study skills. Standardized instruments were administered to treatment and control students and appropriate statistical analysis procedures were used to examine several research areas including:

- The pattern of student scores during the 1973-1974 school year and the previous school year in terms of their means, standardized scores, and performance in relation to national norms.

- The statistical significance of differences in the mean scores of treatment and comparison school students examined.

- The statistical significance of year to year changes in mean scores among groups of students.

This approach provided a comprehensive profile of treatment and comparison school students and allowed for continuity in the reporting of student cognitive data. In addition, it allowed for an evaluation of the performance of ESP students as compared to students in the comparison schools. Finally, the Level II approach examined the process of treatment and comparison school students through a longitudinal tracking system.

The Tests of Basic Experiences (TOBE) and Comprehensive Tests of Basic Skills (CTBS) were administered in spring 1973 and 1974 to measure student achievement. The TOBE was administered to kindergarten and first grade, and the CTBS to second through twelfth grade. All students in the ESP elementary schools, the ESP middle school, and those in grades 9 and 11 in the ESP high school were required to take the achievement tests. All students in the comparison elementary schools, and those in grades 9 and 11 in the comparison high school were also tested. A total of 6,282 students was tested, which represented 90% of students in the relevant schools or grade levels based on EISD enrollment figures.

Classroom Differences

At the kindergarten level, significant differences were found from the ANOVA among the classroom means of the six schools on the TOBE language and social studies subtests. These differences were located by applying Scheffe's post hoc multiple comparisons, and found to exist between two treatment schools. One of these schools had the highest subtest means, the other had the lowest classroom mean on both subtests among the six schools. The investigators found no information which might explain why the one school had the lowest mean on these subtests. However, there was evidence that the high mean achieved by the students in the other was related to a curriculum directed at both language and social studies developed by one of the kindergarten teachers in that school. This curriculum was designed under the aegis of the ESP program, consisted of bilingual material, and was used to varying extents by the other kindergarten teachers in that school.

At the fifth grade level, the ANOVA procedure revealed a significant difference among the classroom means of the six elementary schools, and among the language subtest means. This time, however, Scheffe's post hoc multiple comparisons revealed no significant differences. Nor was the comparison between the classroom mean of the treatment schools (36.57) and the classroom mean of the comparison schools (35.22) significant using this same method. Recognizing that the ANOVA procedure still revealed a significant difference somewhere among the classroom means of the six schools, two less conservative tests were employed to locate any differences between the schools, namely, the Newman-Keuls and Tukey tests.

These two revealed that the only significant difference on the language subtest at the fifth grade level existed between the highest classroom mean of one school (ESP) and the lowest classroom mean of another (comparison). A program called the "Diagnostic Reading Program" was in effect in all the treatment elementary schools during 1974. Apparently the fifth grade teachers in the ESP school were more successful in implementing that program.

Finally, both the Newman-Keuls and the Tukey procedures were applied to the difference on the fifth grade language subtest between the classroom mean of the treatment schools and that of the comparison schools. Neither test proved the difference to be significant.

The treatment ninth and eleventh graders scored significantly higher than their comparison counterparts. Several treatments were in effect in the ESP high school, including Man: A Course of Study (MACOS), best paced math, college preparatory English, and language development in social studies. The approach to these programs involved teaching the students problem solving skills, improving their writing, and providing them with individualized attention. The combined effects of this approach appeared to have had an impact on all the achievement subtests.

The major findings obtained in comparing the mathematics subtest means were:

- ESP ninth scored significantly higher than the comparison ninth.
- ESP eleventh scored significantly higher than comparison eleventh and ESP ninth.
- Comparison eleventh scored significantly higher than the comparison ninth.
- ESP eleventh scored significantly higher than the comparison ninth.
- Comparison eleventh did not score significantly higher than ESP ninth.

These findings suggested that the ESP best paced mathematics program appeared to have had an observable effect on both the 9th and 11th grades in the treatment high school, as evidenced by the relatively higher scores in mathematics.

Year to Year Changes

In order to determine if there were significant differences in changes in student achievement scores between 1973 and 1974, the ANACOVA design was used. The scores of treatment and comparison school students, at the elementary and high school levels, for 1973 and 1974, were examined.

At the elementary level, the 1974 CTBS scores of all treatment sixth graders were compared with the scores of those same students who took the battery in 1973 as fifth graders. These changes were then tested against the changes in the 1974 CTBS scores of comparison sixth graders who were administered the battery as fifth graders in 1973. A similar analysis was performed on the differences between the 1974 CTBS scores of ESP eleventh graders and their 1973 scores as tenth graders, and against score differences among comparison students at the same grade levels.

There were no significant differences between treatment and comparison fifth and sixth grade students in the changes from their 1973 to their 1974 CTBS mean scores in reading, language, study skills, and on the total battery of achievement tests. Moreover, no differences were found between ESP and comparison fifth graders on these subtests in 1973. Finally, no significant differences were found between ESP and comparison sixth graders in 1974. Thus, there was no evidence that the ESP program produced any significant differential effects on the treatment sixth graders.

In regard to the mathematics subtest, however, even though no significant differences were found on this subtest between ESP and comparison fifth graders in 1973, nor between ESP and comparison sixth graders in 1974, the changes in the ESP fifth graders from the first year to the second were significantly higher than those of their comparison counterparts. The ESP students gained 7.58 percentage points from 1973 to 1974, which was significantly higher than the comparison students' gain of 4.95 percentage points. Since the best paced mathematics program was the dominant treatment in effect at the ESP elementary schools in 1974, it appeared to have produced positive differential effects.

No significant differences were found between treatment and comparison students, at the 10th and 11th grade levels, in changes from their 1973 → 1974 mean scores on the reading, language, and study skills subtests. The data reported in the first interim report indicated that there were no significant differences on these same subtests between treatment and comparison 10th graders tested in 1973. However, significant differences were obtained between treatment and comparison 11th graders tested in 1974.

Although there were no significant differences obtained, the comparison 10th graders scored higher on the CTBS subtests than the treatment 10th graders in 1973. From 1973 to 1974 the comparison 10th graders scored lower in the reading, language, and study skills subtests. During this same period the treatment 10th graders gained in their scores. By 1974 the treatment 11th graders scored significantly higher than the comparison 11th graders. This would suggest possible differential effects associated with ESP at the high school level.

The analysis of the ANACOVA results on the math and total battery subtests at the high school level were somewhat different. The data revealed that the changes on the math subtest means of the treatment 10th graders from 1973 → 1974 were significantly higher than the changes in the means of the comparison tenth graders during this period. This was

evidenced by a gain of 8.31 percentage points by the treatment students which was significantly higher than the .47 gain obtained by the comparison students. Since no significant differences were found in the changes from 1973 → 1974 subtest scores, except for mathematics, and since all subtest scores, including math, contributed to the total battery scores, it was posited that the significant year to year increases obtained in the math means caused the increases in the year to year total battery means to be significant.

EVALUATION AND IMPLEMENTATION

The ESP rationale stated that poor student performance was directly related to cultural conflict, which in turn was manifested in student alienation, rejection, mistrust, and low achievement. The affective data collected and analyzed indicated that, with few exceptions, ESP students were not atypical of students elsewhere and had similar attitudinal, self concept, and motivational characteristics. These data suggested that student culture conflict was not as pronounced as originally stated but rather that differences were primarily linguistic and experiential in nature. The treatments that improved student language and mathematics scores spoke to student needs: individualized instruction that allowed for whatever differences existed.

The collective data presented also pointed to the need to reassess the appreciation of comprehensive educational change theory in the field setting. Under idealized circumstances, comprehensive change is composed of education program linkages which speak to student needs and lead to improved student achievement. As operationalized by ESP, comprehensive change was a diverse set of discrete activities poorly integrated into systematic improvement of student achievement. ESP needed to focus on this goal, reassess its operationalization of comprehensive change, establish pedagogical priorities (such as in language and mathematics), and utilize its resources on the establishment of K-12 programs in these areas.

A COMPARATIVE ANALYSIS OF ENGLISH AND SPANISH READING PERFORMANCE OF MEXICAN AMERICAN STUDENTS

The Edgewood School Plan of the Edgewood Independent School District, San Antonio, Texas, had as one of its goals, some instruction and classroom management in Spanish. The district also participated in the state bilingual program and in several special, federally funded bilingual programs. A concern of the ESP Final Plan had been the enhancement of the students' affective, and ultimately, cognitive development through the use of the students' primary language as an integral part of the program.

The Guidance Testing Associates (now owned by St. Mary's University) Test of Reading and Prueba de Lectura, and the reading subtest of the Comprehensive Test of Basic Skills, form Q, were the instrumentation for this study. All three tests were administered within a four week period so the results could be compared.

This study was conducted to ascertain whether bilingual Mexican American students scored significantly different on parallel Spanish-English reading tests and to provide Spanish language test data. Students were selected from two different bilingual programs and on the basis of Spanish language fluency from grade 3 (N=84) and grade 5 (N=76). Additionally, contrasts between these students' scores and those of a grade 8 sample (N=96) on the English GTA and comparable CTBS subtest were examined. Correlation coefficients between the English GTA and CTBS vocabulary and comprehension subtests for grades 3, 5, and 8 were obtained.

Dummy variable regression analysis was used to examine categorical and interval data relationships. The data revealed that the grade 3 and 5 samples obtained significantly higher scores on all English GTA subtests as compared to the Spanish GTA. There were, in addition, several sex and group differences.

The correlations between the English and Spanish GTA subtests were relatively high at the grade 3 level and ranged from .71 to .84. The grade 5 correlations were moderate and ranged from .43 to .68. The correlations between the English GTA and CTBS on comparable subtests for grades 3, 5, and 8 were also moderate. There were no significant differences between these students' standardized English GTA or CTBS subtest scores.

This study raises serious questions about the commonly held assumption that bilingual students' performance on a Spanish test will be superior to that in English. The grades 3 and 5 English and Spanish test data demonstrated that testing exclusively in Spanish would have seriously and detrimentally underestimated the competencies and abilities of these bilingual-bicultural Mexican American students.

Many educators have seriously questioned and challenged the propriety of testing bilingual Mexican Americans in English and have called for "culturally relevant" and Spanish language testing. In many instances, such criticisms are valid and do not excuse the past or present injustices

committed by inappropriate testing and deleterious use of test data. However, as suggested by this study, testing in Spanish may have as many dangers as testing in English. Clearly, inappropriate testing in whatever language can only result in limiting the educational and social opportunities of students. It may well be that multiple measures are essential to more accurately derive needed achievement data.

In addition, the findings also suggested that educational practitioners should closely examine the quality of bilingual programs. If such programs are intended to foster Spanish and English language competencies, as those purported to do in this study, the data suggested they are falling short of their objective. Spanish language fluency of bilingual-bicultural students in a predominantly Mexican American school district (90%), from which students in this study were selected, appeared to decrease considerably after early elementary grades. While administrators and designers of the bilingual programs from which students sampled in this study were selected may have labeled their programs "bilingual," the de facto use of a Model I program design in which Spanish language training received little intensive emphasis beyond grade 2, resulted in a dramatic displacement of Spanish language among Mexican American students within the school setting. A re-examination of these bilingual programs both with respect to their philosophical premises and their pedagogical practices appears warranted.

Lastly, this study suggested that there is a dire need for additional empirical studies in the area of Spanish-English language dynamics with respect to student achievement outcomes. Assessment of language dominance, pre and post English-Spanish testing, and controls of home environmental and school variables would provide much needed information toward understanding the numerous complexities of the educational process as it related to Mexican American students.

AN INSTRUMENTATION STUDY OF THE PURDUE SOCIAL ATTITUDE SCALES FOR PRIMARY CHILDREN (ENGLISH AND SPANISH VERSIONS)

Reliability and validity studies of the English and Spanish versions of the Purdue Social Attitude Scales for Primary Children, Experimental Edition II, were conducted in South Texas in spring 1975. The large samples (English=1620, Spanish=570) included mostly Mexican Americans but some Anglos, Blacks, and Orientals in a test-retest design on each language.

Congruency of item response for students repeating the administrations was very high within languages (82 to 94% for English; 78 to 92% for Spanish) and fairly high between languages (English-Spanish 52-78%). Response frequencies were similar within and between languages. Item discimination between low and high scorers on the subtest and total scores was always in the appropriate direction and the contrasted item mean scores were significantly different. Item Pearson correlations were moderate (.35 to .55) with subtest scores, low to moderate with total scores (.27 to .47). Correlations among subtests were upper moderate (.56 to .63), and high for subtest to total scores (.82 to .84). Factor analysis generally confirmed the subtest structure but indicated some need for improvement. In five corollary studies, students gave reasonable answers for the choices; in sociometric analyses, choices for play correlated highly with the peer subtest scores. English-Spanish intercorrelations were in the lower high range (.78 to .79), indicating viability of the Spanish version.

Correlations between the English and Spanish versions of the UCLA Center for the Study of Education's School Attitude Questionnaire were high (.72 to .91) for kindergarten when compared to the Purdue school subtest scores, adding validity support to the instrument. Correlations for first and second grades were negligible (.11) with the motivation scores of the Self-Concept and Motivation Inventory, low (.23) with the self concept scores of that instrument. Correlations with the language subtest scores of the TOBE and CTBS/Q were low (.37) but high (.85) with teacher assigned grades for language.

The Purdue was easy to administer and score. Children "enjoyed" the instrument, according to teacher and monitor reports. Administration to groups of about ten was recommended.

Kindergarten scored highest in English; there was a drop in first grade and near recovery in third grade. The Spanish scores were similar across grades with third the highest. Blacks and Orientals scored higher than Anglos and Mexican Americans on the English version but the Anglo, Black, and Oriental samples were so small in relation to that of Mexican Americans that ethnic difference conclusions were questionable.

No ethnic group differences appeared in Spanish. Mexican Americans scored significantly different across schools; the other ethnic groups did not.

A tentative stanine chart was produced based on the major English sample of 1620 students. The South Texas pupils scored about like data provided by the instrument author, Dr. V. G. Cicirelli, from Gary, Indiana. The stanines for the samples from the two states were alike.

Recommendations to practitioners included general confidence in the instrument for grades K-3, both in reliability and validity, and in both languages, in measuring the subtest constructs of peer, school, home, and community attitudes and the instrument's overall attitude concept. It apparently did not predict standardized language achievement scores well but it did predict the grades the teachers assigned to the pupils. It did not appear to predict either of the affective constructs in the Self-Concept and Motivation Inventory.

Research was urged on large samples of Orientals and Indians. Anglos and Blacks in areas other than Texas and Indiana should be researched. Spanish speakers in other regions should also receive additional study, especially with the Spanish version since its vocabulary was tailored specifically to South Texas. Additional validity and discrimination studies were also suggested.

SOCIAL ATTITUDES OF SOUTH TEXAS PRIMARY STUDENTS

One of the National Institute of Education's urban experimental schools sites was the Edgewood Achool Plan in San Antonio, Texas. Its predominantly Mexican American student body was served in a locally developed and administered program designed to improve their affective domain, which in turn was expected to enhance their academic achievement. The program's basic document reflected the findings in much of the literature: Mexican American students begin school with negative attitudes and these worsen with exposure to education.

Development Associates, contracted to do the external evaluation of the program, keyed its research to the specific guidelines of the district's documents and studied the attitudes of the students, as well as other factors such as academic achievement, over the first four years of the five that were funded. The students were primarily bilingual, English and Spanish, and the experimental program was concerned that language bias not interfere in the classes or the testing.

The Purdue Social Attitude Scales for Primary Children, Experimental Edition II, was chosen for the attitude study in grades K-3 in the Edgewood School Plan and in the comparison schools. A parallel Spanish version was produced and investigated. It was found to have high reliability and strong validity. The samples were - 1975: English 1620, Spanish 570; 1976: English 711, Spanish 407 - in a one year longitudinal study. These were drawn from the experimental schools and two groups of comparison schools, others from the same district and some from other small town districts in South Texas.

The Edgewood comparison school means were generally higher in English than the experimental but not always so. Both were significantly higher than the small town schools. In Spanish the two Edgewood groups scored alike and significantly higher than the small town schools. Some differences were found among the scores of Mexican Americans, Blacks, Anglos, and Orientals in English but because of the great variation in sample sizes, these results were tenuous. The ethnic groups did not score differently in Spanish. Mexican Americans' scores varied significantly across schools, the other ethnic groups' scores did not. Sex was not a factor in any of the analyses. Students in kindergarten scored highest, followed by third, second, first. These differences generally remained when a sub-sample of the same students tested in 1975 were retested in 1976.

While the scale was not negative-positive polarized, the "very sad" to "very happy" graduated scale could infer such interpretation. No mean would have fallen below the middle point in the scale and most means would have approximated "happy," when calculated by groups of schools, schools, grades, or ethnic groups. This was true for either the English or Spanish version scores. Four subtests made up the total score: peers, school, home, and community. The pattern was for peers to have the highest score and declining slightly through school, home, and community.

The significant differences followed the patterns of the total scores.

The first, second, and third grade subtest and total scores were compared with data from Gary, Indiana, administrations; those for first, third, and all students were not statistically different; the Indiana second grade was higher than the Texas second. Texas and Indiana Blacks scored alike; Texas and Indiana Anglos did also. The Indiana data did not report Mexican Americans but since their scores were essentially like the other two groups, no important difference was found.

The experimental schools funding document had contended that their Mexican American children started school with a negative attitude and that the condition became worse with more time in school. However:

1. The Edgewood kindergarten students in the sample started school with positive attitudes, at or above the norm stanine of 5.
2. Those same kindergarten children, one year later in first grade, had numerically higher mean raw scores and at the same or next higher stanine.
3. Although there was a significantly higher score for kindergarten than first grade in 1975, both were at stanine 5 and the 1975 first graders raised their scores in 1976, as did all grades.

Neither of the two program contentions were upheld from the data collected in this study.

The experimental plan proposed to improve the students' attitudes through several programmatic interventions. However:

1. The scores of the experimental and same district comparison students were usually alike; in a few cases, the comparison students were higher.
2. The experimental students showed gains in scores over the one year period; the comparison showed about equal gains across the year also.

No discernable effect of the experimental program was found from the research data. Further, the likelihood of language bias existing, when applied to the sampled children as a whole, was not upheld from the research. Bilingual children scored about alike in English and Spanish in all grades except kindergarten, where the students scored higher in English, and in the small town schools where all grades scored higher in English.

The results lead to the conclusion that educators need concrete data from valid instruments before they ascribe some negative effective characteristics to Mexican American students and before they build a program based on such characteristics. If Mexican Americans generally are like the students in this study, treating them en masse as negative would be a serious programmatic error.

LOCUS OF CONTROL AND ACHIEVEMENT OF MEXICAN AMERICAN PUPILS

The Edgewood School Plan of the Edgewood Independent School District of San Antonio, Texas, outlined a series of actions it would take to improve the learning environment of its students, including devising strategies that would address the specific needs of its primarily Mexican American student body. Although the thrust of several of the activities it planned to carry out were specifically identified, it was recognized that the scant literature about academic achievement in Mexican Americans made it imperative that some causes of differential learning be investigated. Utilizing the Nowicki-Strickland Abbreviated Elementary and secondary forms and the appropriate levels of the Comprehensive Tests of Basic Skills, form Q, samples of 79 third graders and 59 students from the combined seventh and eighth grade classes in one of the middle schools were studied in spring 1975.

The results of this study may be summarized as follows:

- At grade 3, neither locus orientation nor sex was significantly related to differences in achievement scores.
- Externally oriented 7 & 8 grade students scored significantly higher on reading, arithmetic, and total battery achievement scores than did internals.
- Seventh-eighth grade females scored significantly higher on the language achievement subtest than did males.

It should also be noted that locus orientation did not significantly affect language achievement scores of the 7 & 8 grades sample, even though externals continued to score numerically higher.

If subsequent research supports these findings, it may be that keying instruction to locus orientation will enhance achievement. That is, for students that are externally oriented, a clearly defined and structured instructional program may be appropriate; conversely, for internally oriented students, an individualized self directed program may function better. The implications of this research are that if the academic personality disposition of students could be identified, and the instructional content keyed to those dispositons, both the teaching and learning processes might be facilitated. The present study, in conjunction with the results of similar investigations, appear sufficiently uniform to warrant a feasibility study into gearing an educational program toward locus of control orientation.

LANGUAGE DOMINANCE TEST DEVELOPMENT AND INSTRUCTIONAL PLACEMENT: A CASE STUDY

An important premise of bilingual-bicultural and compensatory educational programs is the need for instruction in the pupils' "best" language. In theory, determination of pupils' "best" or dominant language, followed by appropriate instruction in that language, would produce increased achievement. Although many instruments have been developed, the assessment of language dominance has been problematic. In some cases, school districts have utilized existing instrumentation while others have sought to develop their own.

A case study of the efforts of an experimental program, the Edgewood School Plan of the Edgewood Independent School District, San Antonio, Texas, to develop a language dominance instrument, was conducted to examine the process as part of a comprehensive evaluation of that experimental program and to provide information to practitioners on the instrument developed and general test development. The study consisted of documentation of the development process, a review of the instrument by a panel of experts, examination of relevant program documents, analyses of the instrument's psychometric properties, language classification coincidence with a standardized test, placement of the pupils into instructional language modules, and the possible relationships between the language classifications and achievement.

In summary, the findings of this case study were:

1. No reliability nor validity data were presented by the instrument authors.
2. There appeared to be inadequacies in the process of reliability and validity studies of the items in the instrument.
3. Response categories, number of items in sections, and item construction cast doubt on the face validity and reliability of the instrument.
4. Possibilities of rater bias appeared to be a serious weakness.
5. The coincidence between classifications of the local instrument and those of the James Language Dominance Test was not significantly better than by random classification.
6. Teachers judged the local instrument's placement to be inaccurate and a relatively large number of pupils was moved among the instructional modules.
7. There was no relationship between achievement scores and either of the two instruments' classifications at kindergarten or first grade; the James Language Dominance Test was significantly related to the achievement scores at the second grade with the bilingual students scoring higher than those designated as Spanish dominant.

Language dominance placement and instruction had not resulted in improved achievement of the youngsters in the sample, at least as hypothesized by

the experimental program. Further, even if the "best" language theory had been adequately tested in the classroom instruction, the instrument developed by the experimental program presented such serious deficiencies that its use for this purpose could not be recommended.

The results of this study suggested that educational practitioners should not undertake language dominance instrument development unless they are capable of carrying out the scientific processes necessary for such development and unless they are prepared to commit extraordinary fiscal and other resources to the effort. Finally, the authors of the present research suggested that there was an urgent need to reexamine the premises on which language classification, instructional placement, and achievement testing are based. The Mexican American students in this study appeared to have been further handicapped by the process utilized by this experimental school program.

X

PSYCHOSOCIAL GROWTH AND ACADEMIC ACHIEVEMENT OF MEXICAN AMERICAN STUDENTS

The Edgewood School Plan, a National Institute of Education experimental schools project in the Edgewood Independent School District, San Antonio, Texas, developed a rationale for its program that sought to right what it saw as the "incompatibilities between the predominantly Mexican American student body and a middle class, Anglo Saxon system." It specifically attempted to reverse "negative self concept, motivation, and attitudes," and by so doing, improve the academic achievement of the students.

The program acquired more Mexican American teachers, used Spanish as a language of instruction and classroom management, developed and/or adapted materials with Mexican American and local content, increased small group and individualized instruction. Although considerable variation was present in each of these educational elements during four years of federal funding, the original goals were reaffirmed in their continuation documents.

Development Associates conducted the external, summative evaluation of the project, and carried out on site observations, measured each of the affective and cognitive variables, and studied the curricula and their implementation during the first four years of the operation. The present research represents the aggregation of the affective and cognitive studies across these four years, in order to provide evidence on the premises of the plan and how well it accomplished its goals during the period.

The research design included both cross sectional analyses within each of the years and longitudinal analyses across the four. Students were tracked across years as cohort groups: Cohort I began with the first grade and progressed to fourth; Cohort II was initiated with fifth and continued through seventh; Cohort III covered eighth through tenth. The last years in the latter cohort groups suffered considerably from attrition. Both experimental and comparison students were studied simultaneously. The affective domain was measured via the Self-Concept and Motivation Inventory and the Student Behavior Description Questionnaire; achievement was tested with the TOBE-CTBS/Q series.

No major differences were found between the treatment and comparison groups at the beginning of the four year period. Research question one examined the district premise that their youngsters were low or negative in affect. Examination of self concept and attitudes found these Mexican American students at or above the norm for the instruments. They were somewhat below the norm in motivation for some of the years in which they were measured. Achievement was below the norm every year.

Research question two probed the efficacy of the treatment program in improving the affect of the students and thereby raising their achievement scores. The data from all students completing the affective and cognitive instruments in any one year were analyzed cross sectionally; those students present across the tracked period, the longitudinal subsample, were studied cross sectionally and via repeated measures analysis. Blocking for grade,

sex, and group (treatment and comparison), the affective variables (self concept, motivation, and achievement, affiliation, and inclusion attitude) and the cognitive variables (reading, language, arithmetic, total achievement, and study skills) were analyzed where applicable.

Few significant effects were found. Those that did occur tended to be cognitive rather than affective, and most of the differences were between males and females with the latter scoring higher. At grade 4 in Cohort I, the comparison students scored higher in achievement than the treatment students. Fewer differences were found with the longitudinal than with the cross sectional samples. When the cohort groups were studied across time via the repeated measures analyses, no affective differences remained, that is to say, the differences had been short run and had not persisted across the life of the project. One cognitive difference continued: the comparison scored higher in language than the treatment in Cohort I. In summary, the authors concluded that the treatment program had not been efficacious in promoting either the psychosocial or cognitive growth of the students.

Because of the treatment program's proposition that improvement in affect would concomitantly raise achievement, research question three examined the influence of the affective variables on cognitive performance. Utilizing multiple regression analyses with stepwise inclusion and canonical correlation analyses, the amount of variance in achievement explained by the affective variables and sex was determined by grade and sex for the longitudinal subsamples. The few sporadic and inconsistent findings, perhaps due to chance, led the authors to conclude that the effect of affect on the achievement of these Mexican American children was negigible. Sex differences were likewise unimportant in these analyses; grade did not produce significant differences.

The conclusions for practitioners included:

1. In general, the assumption that Mexican American students' affect is low or negative was not upheld; these students were near or at the instruments' norms.
2. There was little or no relationship between the affective and cognitive scores of these students.
3. A large individualized instruction program appeared to have produced a relative decrease in the treatment students' achievement scores when compared to those in a more traditional program at grade 4.
4. Females scored higher than males on a number of variables between third and sixth grades; there were few differences before or after this period.
5. Basing an instructional program on the theory of incompatibilities, at least as translated into action by this experimental program, did not produce improved achievement scores.

Overall, the treatment program of the Edgewood School Plan did not improve the life chances of its Mexican American students through increases in either the affective or cognitive domain scores when these were compared to those of students in similar schools who did not participate in the experimental program.

Attention was directed to three areas of research meriting further study:

1. Why females scored quite consistently higher than males during the later elementary grades.
2. What variables, such as classroom cognitive climate, for example, account for differentials in achievement?
3. Why these Mexican American students apparently achieved more in the more traditional instructional approach than in individualized instruction, at least over the short run in fourth grade.

Answers to these problematic areas should greatly assist schools in designing and carrying out programs of instruction that would measurably increase academic achievement and psychosocial growth.

ACADEMIC ACHIEVEMENT OF MEXICAN AMERICAN STUDENTS: THE EDGEWOOD SCHOOL PLAN

Academic achievement was the ultimate goal of the Edgewood School Plan, a National Institute of Education experimental schools program in the Edgewood Independent School District in San Antonio, Texas. Locally developed and administered intervention programs were implemented for the largely Mexican American student body in four elementary, one middle, and one secondary schools designated to receive the experimental treatment. These schools were chosen from the twenty-six in the district on the basis of constituency, a mini-system of feeder schools K-12. Two elementary, one middle, and one secondary schools were selected as comparison schools on the basis of similar ethnic group composition, socioeconomic level, achievement, and affective characteristics as the treatment schools. Five year federal funding was provided for the project and Development Associates contracted to do the external evaluation of the first four of the five years.

Each year the Tests of Basic Experiences were used as measures of achievement in grades K and 1; the appropriate levels of the Comprehensive Tests of Basic Skills, form Q, were administered to grades 2-12. In 1971-1972, the baseline year, the experimental and comparison groups scored alike except on one subtest. No important differences existed between these groups.

In the second year, the treatment kindergarten and first grade children scored significantly higher on all the subtests than did those in the comparison schools. Fifth graders in the experimental schools scored higher in language achievement than did comparison children. The treatment secondary students scored significantly higher on all subtests and the total battery than those in the comparison high school. A large increase in mathematics accounted for most of the total battery score difference. The gains in mathematics from the first to the second year were greater for experimental than comparison. During this year treatment programs had been in effect in one of the experimental schools' K and first grades, in the language instruction in their fifth grades, in the secondary mathematics and in most of the other ninth grade classes. The differences between experimental and comparison were not uniform across grades or schools except as specified, however.

In the third year, the experimental kindergarten and first grades continued to be significantly higher in achievement than comparison. For grades 2 through 7, however, the only significant differences were in favor of the comparison students. Experimental grades 8 through 12 continued to exhibit higher scores in some subtests than comparison. The research design examined the role of sex factors. Several of the significant differences were accompanied by significantly higher scores for females, especially those in the experimental program. An examination of the treatment curricula gave no indications that the thrust of the experimental programs favored one sex over another.

The fourth year achievement data showed the experimental kindergarten and first grade students continuing higher in mathematics than comparison students but not in language. Treatment second graders were higher on all subtests and the total battery. There were no differences at third grade between the two groups. There were several statistically different scores at fourth through sixth grades, all in favor of the comparison students. Seventh grade data showed no differences by group. Isolated differences in favor of comparison occurred at the eighth and ninth grades. Scattered differences with experimental higher were found at the other grades. Although fewer than for the previous year, a number of sex differences, all with females higher, remained during the fourth year, 1975-1976. In many cases throughout the K-12 span, the treatment students scored lower during the fourth year than the corresponding grades had scored the year before. The several elementary schools still showed significantly different means, with one comparison school highest and some treatment schools decidedly lower than this school. An examination of the programs in these schools did not provide a consistent explanation for the differences.

Several of the data patterns evident in previous years were interrupted in 1975-1976. Some of the statistical differences occurred, or were not repeated, because the experimental students' scores were lower than the previous year, rather than that the comparison scores had changed. Relatively massive individualized instruction treatments were infused into most classes in the experimental schools in the fourth year. The comparison high school also introduced substantial individualized curriculum packets into its program. Teachers and students may have experienced adjustment difficulties with these efforts that detracted, at least temporarily, from the general progress.

Examination of the cross sectional achievement data across the four years provided important general information:

1. Neither the treatment nor the comparison group changed noticeably over the four years from the pattern of becomming increasingly further behind the national norm as the students progressed through grades K-12.
2. Performance in the elementary schools was not uniform; neither the experimental nor the regular instructional program produced consistent results across schools.
3. The achievement means of the treatment and comparison groups did not remain stable across the four years; the comparison group showed some losses the third year, experimental the fourth.
4. Performance was not uniform between girls and boys; the experimental females were frequently and often substantially higher than experimental males and either sex in the comparison schools.
5. No convincingly consistent pattern of achievement differences was encountered when the scores were analyzed by performance quartiles or when compared to national norms.

When these generally inconsistent patterns were taken into account, the conclusion must be that the experimental program was unable to produce general academic achievement increases with its treatments.

In brief, the Edgewood School Plan, during its first four years of funding, introduced a variety of treatments over different grades and for varying lengths of time. Some consistency of treatment was observed in the general approach and in the staff training supplied to teachers. When viewed annually, some signficantly higher experimental students' scores appeared to be associated with treatments in the particular subjects and grades where differences between experimental and comparison groups were noted. The experimental program, however, tended to change its treatments often or at least to modify them considerably; persistence of treatment effect was not encountered. In the aggregate, the considerable inconsistencies led to the general conclusion that the experimental program had not substantially improved the achievement scores of its pupils when viewed by school, grade, or subject matter in relation to the Edgewood comparison students that did not participate in the program.

AN INSTRUMENTATION STUDY OF THE SELF-CONCEPT AND MOTIVATION INVENTORY (English and Spanish Versions)

The Edgewood School Plan, an experimental schools program of the National Institute of Education in the Edgewood Independent School District, San Antonio, Texas, emphasized the enhancement of the students' affective characteristics as means toward improving their academic achievement. Self concept and motivation were important elements in this plan. The external evaluators of the program, Development Associates, in consultation with its technical panel and district officials, chose the Self-Concept and Motivation Inventory (SCAMIn) as the most viable of the available instruments for the measurement of these two constructs. The inventory was administered to grades K-12 samples during four of the five years of federal project funding.

To examine the possibility of language bias in the testing, Spanish versions of the four forms (preschool, early elementary, later elementary, and secondary) were produced and substantially pretested. In spring 1975, students were tested three times: English, Spanish, Spanish, each approximately three weeks apart. Interviews were conducted with from 30 to 50% of the tested students druing the two weeks following the last administration. All the children were bilingual Mexican Americans and the sample sizes for the testing were: kindergarten=55, early elementary (second grade)=78, later elementary (fifth grade)=63, secondary (eighth grade)=69. These samples' mean scores were compared to those of other administrations and years and no substantial differences were noted, indicating general representativeness of the groups from which they were selected.

The Kindergarten Study

Viewing English-Spanish as a test-retest showed fairly high congruence of reply and mostly acceptable item intercorrelations for the motivation test across administrations. Some validity questions were raised on some items, nevertheless, that correlated at a low or negative level with the motivation score. The Spanish version performed approximately the same as the English but some items in both languages needed improvement.

The self concept portion of the SCAMIn presented problems in both reliability and validity for this sample of Mexican American children. An examination of the internal consistency appeared satisfactory in all three administrations and the subtest and total means did not vary in unusual amounts. The item congruencies were low, however, and the item intercorrelations were widely differing with many too low. Interviews with students showed poor comprehension of several items and some problems with abstractions and unusual wording. The Spanish version performed slightly better than the English but both needed substantial improvement.

In comparisons between the two major constructs, motivation and self concept, and their subtests, considerable overlapping was found between

self concept and the motivation subtest of achievement needs. The other motivation subtest, achievement investment, did not exhibit that same property, demonstrating a great deal of independence.

One of the major difficulties noted for these applications of the SCAMIn was the almost uniformly high item means. These appeared to be due to a general tendency to respond positively to the items, rather than a response set toward the right hand alternative on the answer sheets, since the achievement investment subtest, which involved reverse scoring of the items, did not have notably lower means than the other subtests. The discrimination powers of the instrument to detect changes across time appeared to be questionable in its present form.

The Early Elementary Study

This study, as it functioned with the sample of second grade Mexican Americans, showed low correlations between the two subtests of motivation, achievement needs and achievement investment. The achievement needs mean scores were similar in the three applications. The achievement investment and the total motivation scores differed significantly between the English and one of the Spanish administrations but not between the others. Internal consistency reliabilities (Cronbach's alpha) were moderate but acceptable for motivation and achievement investment, but low for achievement needs.

While the incongruence of reply across administrations was high, suggesting poor reliability, the interviews appeared to show that much of the item response change was due to situational differences that had occurred during the intervening period. Seen, then, as capturing these changes, the motivation section generally performed as would be expected. Three items suffered comprehension problems. The Spanish version functioned about like the English.

The self concept analyses found moderate relationships between the two subtests, role expectation and self adequacy. The subtest and total scores correlated mostly moderately across administrations.

Eight of the twelve self concept items had apparently probed situational changes due to circumstances inherent in the period of time in which the administrations were given. While the item congruency calculations and the intra-item correlations appeared to show poor reliability, they were sensitive, the interviews revealed, to some changes in the pupil's feelings. However, internal consistency reliabilities were unacceptably low for self concept and its two subtests.

As a generality, items in the early elementary form of the SCAMIn performed more appropriately than had items in the kindergarten form. Some of the same difficulties with items containing abstract time references occurred in both. Item comprehension was more uniform and more accurate for the second graders on the early elementary form than for the kindergarten students on the kindergarten form.

However, the poor reliability indications called the self concept portion of the early elementary form into some question. The SCAMIn authors furnished only a single test-retest reliability of .77 for the total instrument over a two week period. The test-retest data and internal consistency reliabilities from the present study suggested that the reliability of the self concept portion of the SCAMIn with this Mexican American sample was much weaker than that.

The Later Elementary Form

The responses of fifth grade children to the later elementary form resulted in better psychometric properties than the previous two. While some skewness and ceiling effect was evident in the motivation and its subtest scores, distributions for self concept and its subtests were more nearly normal.

Correlations between the two motivation subscores were quite low, while correlations between the self concept subscores were high-moderate and indicated probably overlap. Achievement needs and achievement investment correlated more strongly with the self concept subscores than with each other, suggesting that the four scores may not be as factorially independent as the SCAMIn authors claim.

Item analyses showed most item-subtest correlations to be moderate. While there were some items with low correlations across administrations, most showed moderate stability. Item congruency was not very consistent across applications. Test-retest score correlations across administrations were moderate to strong, indicating substantial stability, especially between the two Spanish tests. Internal consistency reliabilities were quite good for group use.

Interviews with respondents indicated that the items were understood well both in English and in Spanish. Most items functioned well in both languages although there were a few items which had different meanings in the two languages. Some items were found to have been sensitive to situational variations; others had alternative activities or objects which led to differing responses across administrations. Overall, the psychometric properties of the English and Spanish forms were quite similar. The larger number of items in this form, as well as the increased age of the children, probably contributed to the improved functioning of this later elementary form as compared to the preschool and early elementary forms.

The Secondary Study

Reasonably good psychometric properties were found with this form as it functioned with an eighth grade sample. The distributions of scores for motivation and its subtests showed negative skewness and ceiling effects, particularly for the achievement investment subtest. Self concept and its subtest more nearly approached a normal distribution. The few

significant differences in score means over the three administrations did not indicate any particular trends.

There was substantial overlap between the subtests making up the motivation score and overlap was even greater for self concept. Additionally, achievement needs, and to some extent the total motivation score, correlated moderately with self concept and its subtests. This suggested that the subtests of the SCAMIn are not independent.

Item-total correlations were in the moderate range although a few extremely low correlations were found in self concept. Item correlations across administrations showed good stability between the two Spanish administrations but much weaker for the English to Spanish comparisons. Item congruence revealed the same lack of consistency from English to Spanish. Test-retest correlations showed excellent stability between the Spanish applications. Internal consistency reliabilities (Cronbach's alpha) were satisfactory for group use of the research instrument.

The interviews indicated that items were understood quite well in English and in Spanish and the students had responded appropriately to the intent of the questions. Certain items, however, caused difficulty, particularly when loading and unusual phrasing were present.

Overview of the Four Forms

The instrumentation study had several limitations, including the sample sizes, number of administrations, and teacher judgment on bilinguality rather than instrument measurement. Similarity of scores tended to show, however, that the samples were not substantially deviant from the larger groups. Factor analyses would have added to the results, as would in-depth validation studies.

Although relatively good congruence of reply and high correlation across the subtests and administrations were found, some differences in means and standard deviations resulted, but these presented no particular pattern across the four forms. Throughout, the motivation subtests were less strongly correlated than the self concept subtests. Further, achievement needs tended to have significant correlations with the self concept subtests, indicating overlap among them. This was not sufficiently strong, however, to invalidate the use of the subscores.

The two language versions performed about equally but there were some items that needed improvement if they were to be considered parallel. Sex did not differentiate in any but a few and these were not consistent in pattern.

Indepth interviewing provided evidence for item content validity as well as language congruence and the differences between them. There was some indication that the items functioned better in Spanish than English on the kindergarten and early elementary studies but the reverse was true for the later elementary and secondary studies.

A number of items reflected short term changes in the children's situations; this was more typical of the kindergarten and early elementary forms. Items with more than one affective object, or more than one action alternative, also presented problems. Some items had words or phrases that were not well understood; again, this was more frequent in the first two forms.

Further studies of the validity of the SCAMIn with Mexican American children need to be carried out. This is particularly important in the case of the Spanish version, which needs more investigation. Considerations of content validity would lead to the belief that the Spanish version is about as valid as the English but the fluency of children in either Spanish or English should be carefully controlled in choosing which language version to administer. Finally, the results of this preliminary research appeared to indicate that the versions of the SCAMIn are satisfactory for group research but care should be taken before using them for individual counseling.

THE EXPLANATION OF EDUCATIONAL CHANGE: CASE STUDIES IN THE EDGEWOOD SCHOOL PLAN

The study of educational change, and concomitantly the process of change, was an important element in the external evaluation of the experimental schools project of the National Institute of Education in the Edgewood Independent School District, in San Antonio, Texas. The project, called the Edgewood School Plan, instituted several treatments into its instructional program. Three of these were selected for study in this investigation. Best paced mathematics, a sequenced individualized program primarily for grades seven through twelve, persisted through years two and three, and in modified form in year four, of the five years of federal funding. Cultural advocacy, begun with the inception of the project in 1971-1972, was intended to interject elements relevant to the Mexican American and South Texas experience into the curricula. Its activities diminished considerably during the third year and appeared to be nearly nonexistent the fourth year. The external high school program set up an individualized and largely independent study arrangement whereby adults who had not completed high school could learn the needed material through the Adult Performance Level program units and receive a regular high school diploma. It began operation in fall 1975, continued through the fifth experimental program year, and was officially adopted as an ongoing district program by the board in spring 1977.

The major thrust of this study of three Edgewood School Plan treatments, each with varying histories in the life of the project, was the utilization of a conceptual model to understand and predict the factors affecting change in this experimental schools project. Questionnaires seeking the perceptions of management (board, administration, Edgewood School Plan staff, principals) and operations (teachers) about the three treatments were completed in spring 1976.

The perspective of the study was that educational change would have been deemed to occur if a particular treatment, financed as a part of the project with federal "seed money," was likely to be incorporated into the district, that is, would receive support for continuance even after such federal funds were no longer available. Four factors, via perceptions, were investigated in each of the three Edgewood School Plan treatments selected for study. These factors were: the conditions indicative of a need or readiness for change, involvement of system actors in the decision making process affecting the change, effects of the treatment with respect to its intended outcomes, and the perceived support reported for its continuance. Each was investigated from the viewpoint of two sets of actors within the school system: administrators (management) and teachers (operations).

Index scores for each explanatory factor were utilized in a multivariate regression analysis with stepwise inclusion which yielded information with respect to the amount of variance in the criterion variable perceived support, explained by conditions, decisions, and effects. Regression analyses were performed separately for each treatment, for system actors as a whole, and for the subsamples of management and operations considered separately. These were intended to provide information about the validity of the model for understanding the change process, to identify those variables in the model that were most useful in both understanding and predicting

change, and to supply information about differences in the perspectives of management and teachers with respect to these variables, and of the importance of such differences to the change process. The number of responses from management was quite small relative to the number for operations.

Analyses of the best paced mathematics treatment indicated that regardless of which group of system actors was used as a referent - all actors, management only, or teachers only - the single variable that accounted for most of the explained variance in support was conditions. For the sample of actors as a whole, conditions alone explained 53% of the variance in support. Effects explained only 6% of the variance and decisions only 1%. When controlling for position in the status hierarchy, the relative contribution of decisions and effects to variance in support showed a reversal for the two groups, management and operations.

The external high school treatment results revealed some important similarities and differences. On the analyses using the whole group of actors, and operations alone, as referents, conditions once again appeared as the single most important variable in the model. For the entire group considered as a whole, and for the operations group, conditions explained 55% of the variance in support. For these same groups, effects explained an additional 7% and 9% of the variance, respectively, while decisions accounted for a negligible additional amount of the explained variance. Since the teachers' responses represented over 95% of the total, the total group and operations results are usually parallel. For management alone, while conditions continued to explain a large proportion of the variance in support (66%), decisions became an important explanatory variable accounting for an additional 15% of the explained variance.

For the cultural advocacy treatment, results of the relative importance of the variables for explaining and predicting support across groups closely resembled those found for best paced mathematics. For system actors considered separately, conditions was the single most important variable, while the importance of the two remaining variables, decisions and effects, was reversed when considering management and labor separately.

These results considered collectively suggested that when contemplating change in an educational system as a whole, conditions and effects are more important factors than involvement in decision making, and that of these two variables, conditions is by far the more important factor. Differences did emerge between management and operations, however, with respect to the perceived relative involvement in decision making and effects of the treatment. Participation in decision making was consistently more important to management than was program effect, while the reverse was true for teachers.

One year later, and unknown to the researchers or the questionnaire respondents, the external high school project was officially adopted by the district board for the year following the termination of the experimental program. The best paced mathematics treatment was at least partially incorporated into the same year's academic program. Cultural advocacy, at least as a distinct treatment, seemed to have disappeared. The likelihood of these three occurrences was predicted from the investigation.

The research team concluded from the findings of the present study, combined with those of others based on similar models, that:

- change can be explained/predicted through an adequate study of the perceptions of a system's actors some time before adoption/rejection occurs.
- the model adapted for the present research is powerful enough in its discriminatory qualities to merit consideration for the study of change.

Some additional studies into the psychometric properties of the instrumentation were suggested, especially factor analysis, to determine the necessity of the effects variable, or item analyses, to strengthen this same explanatory factor. The model and its accompanying instrumentation functioned well, however, and subsequent events related to the three treatments gave credence to this approach to the study of educational change.

TEACHER CLASSROOM BEHAVIORS

The experimental schools project of the Edgewood Independent School District, San Antonio, Texas, designed and implemented a special program under the auspices of the National Institute of Education to improve the academic achievement of its predominantly Mexican American student body. The project, called the Edgewood School Plan (ESP), attempted to improve the affective and cognitive environment of the students through encouraging the use of Spanish, adding culturally relevant materials to the curriculum, and enhancing pupil/teacher relationships. This latter included both the instructional techniques and classroom management.

The present report is one of a series produced under contract by the external evaluation team of Development Associates, Inc., as part of its research into the several aspects and results of the program. Specifically, the instrumentation utilized were the 2a version of the Observation Schedule and Record (OScAR 2a) for the first two years of the four year program, then a revision of portions of OScAR 2a (with project specific additions), termed the Modified Observation Schedule and Record (M/OScAR), for the last two years. The study paralleled other research for the first two years and employed very small samples of experimental and comparison students. Subsequently, the scope of the study was changed to more fully document possible changes produced by the program and the sample sizes were greatly enlarged. Classes were monitored for seven semesters by trained observers from spring 1973 through spring 1976.

The major findings for each dimension of the OScAR instruments included the following:

Emotional Climate

- The emotional climate measure scores were mostly at low, positive levels; few negative scores resulted.
- The experimental and comparison classrooms, as groups, were not different.
- The experimental and comparison elementary classrooms exhibited a nonsignificant parallel trend of increasing scores across the seven semesters; the middle and high school level scores did not show this trend.

Teacher Interpersonal Style

- The classroom scores were predominantly at low, positive levels; some negative scores resulted.
- The experimental and comparison groups of classrooms were not different.
- The experimental and comparison elementary scores showed a nearly parallel, nonsignificant increase from fall 1975 through spring 1976; the middle and secondary schools did not show this trend.

Language Arts Skills Domain

- Instructional emphasis on listening, reading, and verbal skills in the elementary schools and listening, reading, and writing skills in the middle and high schools occupied most of the classroom time.
- The areas of verbal skills, such as oral expression, was a distinctly minor component in the middle and high schools.

Although the experimental and comparison classes were generally alike through most of the observations, in the last set, spring 1976, ESP was found to have moved toward individualization on some measures.

In total, the thrust of the experimental program in this study was to improve the learning environment through:

- Orientation and inservice training for teachers that would enhance the teacher/pupil relationships;
- Provision of culturally relevant materials;
- Encouraging the use of Spanish to aid the communications process.

The implementation of these factors in this program neither significantly changed the emotional climate nor teacher interpersonal style scores over the four year time span, nor did it significantly raise the scores of these measures above those found in the group of comparison classrooms. The findings led to the tentative conclusion that these factors and/or the degree of their implementation were insufficient to cause major modifications in the academic climate, as defined by the dimensions measured.

To address the specific needs of the students, individualization of instruction was a concern of the program. To accomplish this end, the staff and consultants:

- Conducted teacher orientation and inservice training about individualized instruction throughout the life of the program; and
- Designed the experimental program materials for group or whole class instruction for the first three years, with the exception of materials for middle and high school mathematics.

The data collection for the first three years showed few practices related to individualization in the classes. Through that period of time, no important differences were found between the experimental and comparison classrooms. Orientation and training appeared insufficient to cause major shifts away from subgroup and whole class approaches, and toward individualization instruction.

In the fourth project year, the program provided substantial amounts of materials designed for individualization. Seven months later, several factors involved in individualization were observed in ESP and comparison classrooms and these were significantly greater in ESP classrooms than in comparison classrooms for the middle and high schools. Because of the existence of differences at one single observation only, the last in the series, the researchers tentatively concluded that training, enhanced by the appropriate materials, can bring about changes in instructional approaches.

Finally, some research literature has reported improved emotional climate as a result of increased openness in education, primarily through individualization of instruction, using e students' home language, and providing culturally relevant learning materials. Viewing all the results of the present study together, the general conclusion was that, at least in the span of time these aspects of instruction existed in this experimental program, these factors did not substantially enhance the emotional climate in the classrooms.

THE COMMUNICATIONS SKILLS PROGRAM IN THE EDGEWOOD SCHOOL PLAN: A CASE STUDY

The Edgewood School Plan initiated a new thrust in fall 1974 called the communications skills program in five of its six schools in the Edgewood Independent School District in San Antonio, Texas. This experimental program, under the auspices of the National Institute of Education, was a part of a locally designed and implemented effort to improve the academic achievement of its largely Mexican American student body. The teachers received some training on the materials provided for six to eight weeks of implementation.

The research utilized questionnaires to probe the perceptions of the experimental schools staff members, principals, and teachers. Experimental teachers and a stratified random sample of comparison teachers also completed sections of the Dimensions of Schooling Questionnaire to provide data on their perceptions of classroom practices.

The investigation was conducted from the fourth to the sixth week of school, coinciding with the implementation of the communications skills program. The external evaluation team also carried out classroom observations via the Modified Observation Schedule and Record at the same time. A spring 1975 application of the research instruments was planned. Due to the cessation of materials production, the program was discontinued. The questionnaires were thus not reapplied as planned.

The goals for the new thrust, the communications skills program, were perceived by the experimental schools staff as: to improve the education of the students and to do so by utilizing the students' home language (Spanish), make the environment more positive, and make the education of these pupils more relevant to their needs. Bilingualism, through the use of English and Spanish in the instruction and classroom management and by teaching Spanish reading, was to enhance the students' self concept. The staff perceived a generally positive climate for the introduction of the new program, rating themselves, the internal evaluation team, and the principals as highly committed to and able to carry out the activities; the board, administration, teachers, and students were perceived as somewhat less committed and able. A diverse set of opinions about the parents was manifested.

The principals of the five implementing schools perceived the goals about like the experimental staff but placed primary emphasis on the improvement of reading and generally improved student learning. Bilingualism was of secondary importance to them. Their concerns about the program tended to be managerial: timely provision of materials, teacher training, and measurement of program results. Their ratings of the commitment and ability of the reference groups were close to those of the experimental staff members but were slightly more positive for the board, administration, and teachers. They expressed doubts about the amount of parental involvement expected in the program. The climate for the introduction of the communications skills program was seen as generally positive by the experimental schools principals.

The teachers charged with implementing the new program in the experimental schools perceived the goals as had their principals. Their concerns for

the program, too, nearly paralleled those of the principals: quality, quantity, and timing of the curricular materials for implementation. The teachers' ratings of the reference groups were almost exactly the same as those given by the principals with the exception that the rated their principals the highest of any group on commitment and on ability to carry out the program.

Modification of the classroom climate toward openness, including individualization of instruction, was expected as one result of the program. The experimental teachers saw themselves as moving toward individualization and open education practices. The comparison teachers saw themselves in much the same way except in the high school, where they perceived the activities to be more traditional. Both groups noted considerable use of Spanish in the classes.

Findings from a parallel study through classroom observations found the perceptions to be directionally congruent with the observation results. The degree of openness, however, was observed to be less than perceived at the middle and high schools and the amount of Spanish spoken in the experimental classes to be less than the perceived amount. No important difference in language usage was found between the treatment and comparison classes.

Implementation of the new program was not as widespread as indicated by the staff; about 33% of the 49 teachers was utilizing the materials as the primary vehicle of instruction. Fourteen percent did not use them at all. The principals' concerns for the teacher training needed for implementation, plus their and the teachers' concerns about the provision of the materials, appeared to have influenced the amount of implementation.

The use of Spanish in the classrooms, as expected from the program design, did not appear to have increased and was no different from the amount observed in comparison classes. Teachers perceived their classroom practices to be moving toward openness, particularly those related to individualization. Some differences were found between the experimental and comparison groups' perceptions but these were not borne out at the elementary and middle school levels; observations showed the two groups' practices to be about the same. Some difference was found at the secondary level, both in perceptions and in observations; the experimental groups' practices were somewhat less traditional than were the comparisons'.

Thus although the experimental staff members, principals, and teachers had generally perceived a favorable climate for the introduction of the communications skills program, implementation was considerably less than expected by the staff. These findings suggested that managerial concerns, both from the principals' and teachers' standpoint, may be crucial in effecting implementation. Their previous experiences with partial curriculum, combined with possession of only a few weeks of materials, appeared to mitigate against complete adoption and implementation in this case.

SUMMARY REVIEW OF REVENUES AND EXPENDITURES FOR THE EDGEWOOD INDEPENDENT SCHOOL DISTRICT AND THE EDGEWOOD SCHOOL PLAN

The Edgewood School Plan (ESP) was a part of the Experimental Schools Program of the National Institute of Education. In June 1971, the Edgewood Independent School District received an initial planning grant to develop the experimental program. In June 1972, an operational/ developmental grant was received by the district. The developmental phase of the grant, for May through August 1972, was to prepare for the implementation of the program in the 1972-1973 school year. The operational phase of the grant was received to fund the program for 30 months, May 1972 to November 1974.

The program spent 55 percent of the budgeted funds for the first program year and 51 percent the second, 1973-1974. The total expenditures for these two years represented less than half of the funds allocated by the operational grant and the National Institute of Education allowed the program to operate on the remaining monies for two more program years. The 1974-1975 and 1975-1976 budgets were between $350,000 and $400,000 less than the budgets for the first two program years. The program underran the 1974-1975 budget by 23 percent and the 1975-1976 budget by 13 percent. Funds from the operational grant were still unspent after 1975-1976 and the National Institute of Education permitted the program to continue operations into the fifth year, 1976-1977. Thus, funds that were designed to permit the program to operate for two and a half years were used for four full years of operations plus a portion of the fifth year, 1976-1977.

The elementary schools' component of the program consistently spent the greatest amount of the funds in each of the last three years. They accounted for over 20 percent of the expenditures from 1973-1974 through 1975-1976. The high school, curriculum development, and pupil appraisal generally accounted for the next largest proportions of the total expenditures. The program components spending the least amount of the monies were cultural advocacy and supportive instruction services, merged later into other components. These accounted for as little as one percent of the total expenditures for a program year. In general, at least 70 percent, and for three of the four completed program years, more than 80 percent of the funds spent by the program were for payroll costs.

Despite the difficult financial and social conditions in the district, and the ambitious goals of the experimental program, the project did not spend its funds. While it was not possible to isolate a single direct cause or reason for the program's inability to spend the funds, general fiscal conservatism and perceived spending restrictions were major causes given by project personnel.

CURRICULUM AND INSTRUCTION IN THE EDGEWOOD SCHOOL PLAN

The Edgewood Independent School District in San Antonio, Texas, under the US Office of Education and later the National Institute of Education, designed and carried out a five year plan intended to modify its programs in ways that would enable its largely Mexican American students and community to profit more from the education offered. Specifically, the experimental schools program under study set out to make the curricula more relevant to the individual needs of the students and to reflect their experiences as Mexican Americans. The plan further proposed to train the teachers to properly implement the new curricula and to make them more responsive to the needs of the pupils. The entire community was defined as the "student body" of the district and special programs to help the residents achieve a better quality of life were to be instituted.

An external evaluation of those efforts was included in the overall experimental schools design. The present study, a part of that evaluation, was to examine the materials and processes of the implementation effort of the Edgewood Experimental Schools Project, later renamed the Edgewood School Plan. The research was mainly process oriented, describing the district's modification efforts and evaluating them, when possible, through expert analysis, teacher and participant perceptions, and relating these efforts to the changes in the affective and cognitive domains where applicable. Comparison groups were used to assess changes in relation to those achieved in other parts of the district.

Curricular Materials

A curriculum development component in the Edgewood School Plan was principally responsible for the production of the materials. A vast quantity was produced and most of it was implemented at least partially somewhere within the experimental schools program. Many of these materials, however, were discrete units; few were revised after initial implementation and/or incorporated into later editions. Two important exceptions to this isolated unit process were found: the mathematics program was almost continuously revised and expanded through the life of the program and the teacher guides produced the last year contained some materials from previous production or revisions thereof.

The results of the research into the curricular materials showed:

- Some improvement was needed in the physical presentation, especially in the size of print for primary pupils, and in illustration.

- Most units contained instructional objectives and these were generally stated in measurable terms; some needed improvement.

- The applicability to the grades and subjects was judged high except that some, particularly Spanish language materials, had measured readability levels too difficult for the students.

- Strong San Antonio and Texas relevancy content was found in about half the units but content on Mexican Americans, Mexicans, and other Latin Americans was rare except in some middle and high school materials.

- Teachers were generally positive about the materials during the first two years, skeptical the third, and were again positive during the fourth year; a concern about the materials being isolated units rather than building a curriculum persisted throughout the period.

Conclusions regarding the curricular materials included the need for a detailed scope and sequence curriculum as a guide to their production and revision; a continuous review process would have aided in the improvement of the materials; and measurement of the results of the use of such materials would have given important directions to the entire production and revision process.

No important differences in attitudes, self concept, and motivation were found between those students utilizing the experimental materials and those using the more traditional adopted texts. Few differences were found in academic achievement; early efforts in an integrated kindergarten program in one school, English in the high school - some differences in mathematics in the second, third, and fourth years - several non-patterned differences in some high school subjects the fourth year. The longitudinal study of both the affective and cognitive domains found no important differences between the experimental and comparison groups' scores when viewed over the four year period studied. Since mathematics was the only area with fairly consistent improvement in achievement, the general conclusion was that the production of high quality materials, subjected to frequent revision, and implemented over a substantial amount of time, apparently offers a greater expectation of cognitive improvement than any quantity of separate, unmodified units, even if the quality of these is high.

Staff Development

The staff development component faced a difficult task since the teacher turnover rate was very high every year. Thus, sequencing the training from introduction to highly complex implementation strategies was hampered. The variability of the curricular materials, and the changes in their approaches and uses, likewise caused substantial changes in the training program.

Nevertheless, a substantial amount of training was offered, from short duration or subject sessions of a few hours to extended and indepth courses. More than 3400 teachers, principals, aides, parents, and other community members were reported to have attended some portions of the

training. These participants generally rated the training as "useful," between the scale points of "undecided" and "very useful."

Since no affective differences were found in the students and few cognitive differences occurred, it would have to be said that the part staff development was to play in this total process, along with the other work involved, provided little evidence of success. This must be tempered by the realization, however, that the changes instituted in the implementation during the years, plus the funding difficulties, made it unlikely that great measurable results could have been obtained. One set of training sequences, that related to individual instruction, although it changed classroom practices little when carried out by itself, appeared to have had an effect when the individualized materials for the implementation were provided - increases in individualization practices were found in several grades.

Instruction

An important aspect of the experimental program was the improved relationships between the teachers and the students. Measurement over the first four years showed no appreciable differences in emotional climate or teacher interpersonal styles between the experimental and comparison classrooms. Similarly, the proposed emphasis on the verbal portions of language arts instruction toward a balanced program of listening, reading, speaking, and writing skills did not result in experimental and comparison group differences. Increases in the use of Spanish in instruction and classroom management, also expected, were not evidenced; during the fourth year the comparison students spoke more Spanish at the high school and no differences were found elsewhere.

The elementary classroom practices varied little between experimental and comparison - both primarily utilized subgroups as the focus of most attention. Some small increases toward individualization were discovered in some measures of the experimental classes. The practices in the middle schools were primarily "class as a whole" oriented but the experimental school was significantly more toward open education on the selection of materials for student use. In 1976, the only year in which significant differences occurred at the high school, the experimental classes were more open than the comparison on attention to individuals and teacher as a resource person rather than a presenter of planned lessons. As a whole, then, experimental and comparison were alike during the third program year but in the fourth year some increases in the open education measures were registered for the experimental group, chiefly in the high school.

Expert teacher observations pointed out some rather serious problems in language arts instruction, especially in the teaching of Spanish. The observations found excellent implementation in middle and high school mathematics, and in middle school social studies. The experts noted a great deal of variation in the quality of the instruction but judged it, overall, as satisfactory or better.

Student perceptions of the learning environment were measured only in the fourth year. Few differences were found between the experimental and comparison groups and these were not consistently in favor of one group or the other.

Non classroom instruction for students in the experimental schools had embraced many activities during the first two and one-half years but were sharply curtailed thereafter. Those that were conducted were judged mostly positively by the participants.

An increase in the proportion of certified teachers and teachers with the masters degree may have been associated with the general thrust of teacher preparation. A type of community college program was developed in cooperation with a local institution of higher learning and many students took beginning courses in the experimental program. Some received credit for prior learning through work experiences.

Training to prepare students to take the General Equivalency Diploma examinations was conducted one year. Studies of the perceptions of those in the experimental and comparison programs showed few important differences between them except for a somewhat higher rating on the quality of instruction in the comparison program, one that had been in existence for many years.

The alternative high school, open only to Edgewood dropouts or residents who had not completed high school, graduated 52 students the first two years and about 200 others were enrolled at the close of the investigation. The program was officially adopted by the Edgewood board of education and was to be continued the year after the experimental program closed.

Non credit training for adults occurred at intervals through the five years and included mostly short term classes or seminars. Few measures of results were attempted but the subjects appeared relevant to the projects and some individuals were later incorporated into several activities of the experimental program.

Conclusions about the implementation of the program closely paralleled those of the curricular materials:

- Frequent changes in the approaches, methods, and techniques appeared to have weakened the possibility of measurable impact on student affective or cognitive scores.
- Changes in implementation appeared difficult to achieve through training alone but seemed viable when training and materials were provided together.
- Few differences were found in implementation strategies or in the quality of instruction when the experimental and comparison groups were viewed across the four year period.

- Although somewhat speculative since they did not persist very long, it appeared that making training available to adults in more accessible sites increased the participation of adults in these.

- Although there were several avenues available for community members to obtain a high school diploma or equivalency certificate, the alternative high school program appeared to have filled a need and was utilized by the community in rather sizeable numbers.

The overall conclusion was that the implementation of the program produced some measurable benefits to the youth and adults in the Edgewood community. It failed to produce results in some critical areas, especially in the affective and cognitive areas, when measured against comparison groups. And corollary to these, it appeared reasonable to conclude that the frequent changes in implementation may have diminished the opportunities for more substantial results.

Recommendations

Practitioners were reminded that the lack of a detailed scope and sequence guide may lead to "disjointed" and tangential efforts in the production of curricular materials and their implementation. Measurement of the results of treatments is imperative if continuous progress is to be registered by a program. Further, general impressions about the functioning of treatments over a short period may be misleading and cause further production or changes when revision may be all that is necessary for increased success.

Researchers were cautioned that longitudinal research on diverse treatments applied to large audiences, especially if the treatments change frequently, calls for extreme care in designing a flexible investigation that can capture both the test score results and the implications from process. It was recommended that both elements be seen as crucial to the measurement of a program. Further, complete separation of external evaluation from the project to be studied was not recommended since important details may thus be lost, making thoroughness more difficult and cost effectiveness lower.

XVIII

FORMATIVE EVALUATION IN THE EDGEWOOD SCHOOL PLAN

The Experimental Schools Program was initiated by the Office of Education in 1970 to determine the effects of a large scale intervention in education accomplished with the aid of long term (five year) federal funding but carried out primarily through a local decision making process. That program was transferred later to the National Institute of Education. In September 1972, an experimental program, first entitled the Edgewood Experimental Schools Program and later the Edgewood School Plan (ESP), was included in this national effort. A part of the design for these programs included both internal and external evaluation. The latter was contracted to a private company, Development Associates, in the early months of the program in the Edgewood Independent School District. The internal evaluation unit was not begun until approximately one year later. One of the tasks of the external evaluation was an examination into the development of internal evaluation in the Edgewood experimental program and this report is limited exclusively to that task.

The external evaluation team, by arrangement with the district and the National Institute of Education, performed both internal and external evaluation activities the first year. The data reported were viewed positively by the district and the experimental program. A district pupil appraisal office and the internal evaluation unit of the experimental program were set up at about the same time. The three offices were working in concert from that point onward. While a testing agreement among them was effected, specific tasks for each one was never explicitly stated, rather, they were implicit in the documents pertaining to each one.

The study of the Edgewood internal evaluation unit, called Level I, by the external evaluation unit, Level II, was conducted from the premise that the descriptions in the Office of Education documents calling for an internal evaluation unit in the experimental schools program, the listing of activities submitted by the district and the internal evaluation unit, and the contents of certain memoranda foom the National Institute of Education, constituted the work objectives of Level I. The Level II effort therefore documented the development process of Level I in the meetings of the district and the experimental program, measured the perceptions of Level I by the clients of internal evaluation, and, via expert examination, studied the products, primarily the Level I reports.

The internal evaluation unit faced a heavy burden of work bec se of its late start. Most of the other ESP components had been in some stage of development and operation for nearly a year. Internal evalution almost immediately carried out a needs assessment survey and the data from that study, along with others available from other sources, were utilized in a replanning process of the experimental program. The planning resulted in a redirection of some program efforts.

Subsequently, Level I carried out a series of investigations: some for the staff development component, including analyses of the component's training;

one on classroom observations via a unique tri-dimensional model; one on the teachers' perceptions of their classroom management; one which administered and analyzed the Mooney Problem Check List for some grades; a series of studies of the experimental and three other principal's functioning; assistance to the family services component in its documentation for acceptance of the alternative high school; a study of student withdrawals; assistance to the language dominance committee in the development and study of an instrument; and ultimately a revision of the instrument. Reports on these, plus quarterly/annual process documents were submitted to the National Institute of Education. About half the reports were tendered Level II for examination.

During the first year of operations, the ESP staff, principals, and teachers, perceived Level I's work positively. The perceptions by the staff development and family services components, three of the six ESP principals, and the National Institute of Education remained positive through the next three years. The curriculum development staff and teachers registered lowered perceptions after the first measurement, fall 1974, and these did not improve thereafter. Cultural advocacy, a separate component for only two of the four years of Level I existence, indicated mixed perceptions. Relationships with the district pupil appraisal office, Level II, and the district administration were described by all of these entities, and by Level I, as strained. The ESP director supported Level I through the four years.

Expert examination of the Level I documents found them of widely varying quality - from weak to excellent. The studies of the principals, the tri-dimensional classroom observations, and to some degree, the study of the staff development training, were judged technically well done. Various problems were encountered with the others: clerical and some technical flaws, conclusions that over reached the data, failure to address the needs of some audiences adequately, and invasion of privacy by including the names of persons studied. The language dominance instrument was described by a panel of test experts as of questionable validity and they pointed out that no reliability data were presented; the reports about the instrument were technically weak.

The author concluded that the Edgewood School Plan, both from the products of Level I and from the other components, had only partially monitored the implementation of its projects and that in some important cases, had not possessed knowledge about whether or not certain objectives had been met. It was further concluded that some of the Level I products had furnished important data about the operation of the Edgewood School Plan but that others had failed to provide much of use to that organization. Positive relationships with some elements within the project had apparently led to worthwhile cooperation and collaboration. Strained relationships with some other elements within ESP, and some outside that project, reduced the probable benefit of the Edgewood School Plan to some of the several intended audiences.

Recommendations to the district from the external evaluation study included a caution to neither accept nor reject the Level I products without careful study since some contained important data and some had limited utility;

further, that the district should in all future internal evaluations insist on clear agreements of the operations of such a unit and monitor its fulfillment of them. It was recommended that other educational practitioners insist on evaluations of all experimental programs, that the objectives of such programs fall within the institution's goals, and that the institution maintain jurisdiction over all programs within its purvue. The research (evaluation) community was urged to insist on written descriptions of duties and responsibilities of internal and external evaluation units and what their precise relationships with other entities were to be. It was also suggested that placing any one unit in the position of giving data but receiving none in return was psychologically unwise and should be avoided if at all possible. Finally, it was recommended that when more than one evaluation unit is involved in the work, an appeals or arbitration process be set up to resolve differences before they endanger the successful fulfillment of their mission.

Retyped

FORMATIVE EVALUATION IN THE EDGEWOOD SCHOOL PLAN

The Experimental Schools Program was initiated by the Office of Education in 1970 to determine the effects of a large scale intervention in education accomplished with the aid of long term (five year) federal funding but carried out primarily through a local decision making process. That program was transferred later to the National Institute of Education. In September 1972, an experimental program, first entitled the Edgewood Experimental Schools Program and later the Edgewood School Plan (ESP), was included in this national effort. A part of the design for these programs included both internal and external evaluation. The latter was contracted to a private company, Development Associates, in the early months of the program in the Edgewood Independent School District. The internal evaluation unit was not begun until approximately one year later. One of the tasks of the external evaluation was an examination into the development of internal evaluation in the Edgewood experimental program and this report is limited exclusively to that task.

The external evaluation team, by arrangement with the district and the National Institute of Education, performed both internal and external evaluation activities the first year. The data reported were viewed positively by the district and the experimental program. A district pupil appraisal office and the internal evaluation unit of the experimental program were set up at about the same time. The three offices were working in concert from that point onward. While a testing agreement among them was effected, specific tasks for each one was never explicitly stated, rather, they were implicit in the documents pertaining to each one.

The study of the Edgewood internal evaluation unit, called Level I, by the external evaluation unit, Level II, was conducted from the premise that the descriptions in the Office of Education documents calling for an internal evaluation unit in the experimental schools program, the listing of activities submitted by the district and the internal evaluation unit, and the contents of certain memoranda from the National Institute of Education, constituted the work objectives of Level I. The Level II effort therefore documented the development process of Level I in the meetings of the district and the experimental program, measured the perceptions of Level I by the clients of internal evaluation, and, via expert examination, studied the products, primarily the Level I reports.

The internal evaluation unit faced a heavy burden of work because of its late start. Most of the other ESP components had been in some stage of development and operation for nearly a year. Internal evaluation almost immediately carried out a needs assessment survey and the data from that study, along with others available from other sources, were utilized in a replanning process of the experimental program. The planning resulted in a redirection of some program efforts.

Subsequently, Level I carried out a series of investigations: some for the staff development component, including analyses of the component's

Retyped

training; one on classroom observations via a unique tri-dimensional model; one on the teachers' perceptions of their classroom management; one which administered and analyzed the Mooney Problem Check List for some grades; a series of studies of the experimental and three other principal's functioning; assistance to the family services component in its documentation for acceptance of the alternative high school; a study of student withdrawals; assistance to the language dominance committee in the development and study of an instrument, and student placement into language modules derived from the instrument; and ultimately a revision of the instrument. Reports on these, plus quarterly/annual process documents were submitted to the National Institute of Education. About half the reports were tendered Level II for examination.

During the first year of operations, the ESP staff, principals, and teachers, perceived Level I's work positively. The perceptions by the staff development and family services components, three of the six ESP principals, and the National Institute of Education remained positive through the next three years. The curriculum development staff and teachers registered lowered perceptions after the first measurement, fall 1974, and these did not improve thereafter. Cultural advocacy, a separate component for only two of the four years of Level I existence, indicated mixed perceptions. Relationships with the district pupil appraisal office, Level II, and the district administration were described by all of these entities, and by Level I, as strained. The ESP director supported Level I through the four years.

Expert examination of the Level I documents found them of widely varying quality - from weak to excellent. The studies of the principals, the tri-dimensional classroom observations, and to some degree, the study of the staff development training, were judged technically well done. Various problems were encountered with the others: clerical and some technical flaws, conclusions that over reached the data, failure to address the needs of some audiences adequately, and invasion of privacy by including the names of persons studied. The language dominance instrument was described by a panel of test experts as of questionable validity and they pointed out that no reliability data were presented; the reports about the instrument were technically weak.

The author concluded that the Edgewood School Plan, both from the products of Level I and from the other components, had only partially monitored the implementation of its projects and that in some important cases, had not possessed knowledge about whether or not certain objectives had been met. It was further concluded that some of the Level I products had furnished important data about the operation of the Edgewood School Plan but that others had failed to provide much of use to that organization. Positive relationships with some elements within the project had apparently led to worthwhile cooperation and collaboration. Strained relationships with some other elements within ESP, and some outside that project, reduced the probable benefit of the Edgewood School Plan to some of the several intended audiences.

Recommendations to the district from the external evaluation study included a caution to neither accept nor reject the Level I products without careful study since some contained important data and some had

Retyped

limited utility; further, that the district should in all future internal evaluations insist on clear agreements of the operations of such a unit and monitor its fulfillment of them. It was recommended that other educational practitioners insist on evaluations of all experimental programs, that the objectives of such programs fall within the institution's goals, and that the institution maintain jurisdiction over all programs within its purvue. The research (evaluation) community was urged to insist on written descriptions of duties and responsibilities of internal and external evaluation units and what their precise relationships with other entities were to be. It was also suggested that placing any one unit in the position of giving data but receiving none in return was psychologically unwise and should be avoided if at all possible. Finally, it was recommended that when more than one evaluation unit is involved in the work, an appeals or arbitration process be set up to resolve differences bevore they endanger the successful fulfillment of their mission.